Investigating the role of Plzf in neural progenitors

Sean Christopher James Constable

March 2015

Division of Developmental Neurobiology

MRC National Institute for Medical Research

The Ridgeway, Mill Hill

London, NW7 1AA

Department of Cell and Developmental Biology

University College London

This thesis is submitted to University College London for the degree of Doctor of Philosophy

Declaration

This PhD project has been completed in the laboratory of Dr David Wilkinson in the Division of Developmental Neurobiology at the MRC National Institute for Medical Research, London.

I, Sean Christopher James Constable, hereby declare that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgements

I would first like to thank my supervisor, David Wilkinson, for giving me the opportunity to carry out my PhD project in his laboratory and for his advice and support. I would also like to thank all the members, past and present of the Wilkinson laboratory for their ideas, encouragement and countless discussions throughout this project.

I am, in particular, indebted to two talented postdoctoral researchers, Javier Terriente and Sebastian Gererty, without whom this project wouldn't have been possible. To Javi, thanks for all the time spent teaching me how to work with zebrafish, for all your advice, both inside and outside of science and the occasional Spanish lessons. To Sebastian, thanks for imparting some of your vast molecular biology knowledge, for your patient explanation of genome engineering concepts and for continued encouragement throughout this project.

Lauren and Rosie, thanks for all the great times whilst students in the lab together and your continued friendship afterwards. Jordi, my thanks both for the many inspiring discussions on genome editing technology and your fantastic tips for places to visit in the world.

My gratitude goes to my thesis committee: François Guillemot, James Briscoe and Alex Gould for their guidance throughout my PhD and insightful comments regarding my project. Thanks also to the aquatics staff who took such great care of my fish. Thanks go to the various people at the NIMR who gave me advice regarding the chromatin immunoprecipitation protocol; Ben Martynoga, Emmanouil Metzakopian and Debbie van den Berg.

Finally I would like to thank my family for their support and backing throughout my studies. Lorna, for being there when the work was tough, for more or less single-handedly planning our wedding and reminding me that there is light at the end of the tunnel.

Abstract

Neurogenesis must be tightly regulated both spatially and temporally to give rise to the full spectrum of cells of the nervous system. The promyelocytic leukaemia zinc finger (Plzf) transcription factor is required for maintenance of stem cells of the spermatogonial and haematopoietic systems and is widely expressed throughout the vertebrate central nervous system. I aimed to understand whether Plzf has a role in the maintenance of neural progenitors, using the zebrafish as my model organism.

To gain information about the transcriptional targets directly regulated by Plzf in zebrafish, I performed chromatin immunoprecipitation on embryos expressing epitope-tagged Plzfa. This analysis failed to identify targets and subsequent troubleshooting determined that the epitope-tagged Plzfa protein wasn't functional.

Analysis of *plzfa* and *plzfb* expression during development reveals that both genes are coexpressed in progenitors in the hindbrain, leading to the hypothesis that the two proteins are functionally redundant. In support of this hypothesis, morpholino-mediated knockdown of the two proteins resulted in a defect in progenitor maintenance. To complement this work, I introduced targeted mutations in the *plzfa* and *plzfb* genes using transcription activator-like effector nuclease (TALEN) technology. In contrast to the morpholino-based results, inactivation of these genes did not result in the same defect in progenitor maintenance; leading to the conclusion that previously unknown off-target effects associated with morpholino use caused the phenotype.

Finally, I developed a simple and highly efficient technique for the insertion of exogenous DNA into targeted locations into the zebrafish genome, aided by the use of TALENs. Using this technique to analyse *plzfa* expression complements earlier analysis suggesting that *plzfa* is expressed in cells actively undergoing neurogenesis. Preliminary functional analysis supports a hypothesis that Plzfa regulates multiple steps during the neurogenic cascade and is important to ensure the correct timing of proneural gene expression during the pathway to differentiation.

Table of Contents

Index of Figures

Index of Tables

Abbreviations

APL = Acute Promyelocytic Leukaemia

bHLH = basic Helix Loop Helix

BMP = Bone Morphogenetic Proteins

bp = base pairs

ChIP = Chromatin Immunoprecipitation

CNS = Central Nervous System

CRISPR = Clustered, Regularly Interspaced, Short Palindromic Repeats

DV = Dorsal View

FGF = Fibroblast Growth Factor

GFAP = Glial Fibrillary Acidic Protein

GFP / RFP = Green / Red Fluorescent Protein

gDNA = genomic DNA

HDAC = Histone Deacetylase

hpf = hours post fertilisation

IHC = Immunohistochemistry

Indels = insertion / deletion mutations

MAB / PBS (T) = Maleic Acid Buffer / Phosphate Buffered Saline (+0.1% Tween)

NHEJ = Non-Homologous End Joining

RARα = Retinoic Acid Receptor α

RFLP = Restriction Fragment Length Polymorphism

RPE = Retinal Pigmented Epithelium

RVD = Repeat Variable Di-residue

ss = somite stage

ss / dsDNA = single stranded / double stranded DNA

SV = Sagittal View

TALEN = Transcriptional Activator-Like Effector Nuclease

TSS = Transcription Start Site

TS = Transverse Section

UTR = Untranslated Region

ZFN = Zinc Finger Nuclease

1 Introduction

The main aim of my PhD work has been to investigate the function of the Promyelocytic Leukemia Zinc Finger (Plzf) protein within neural progenitors of the vertebrate central nervous system (CNS). In order to achieve this aim, I have made use of the zebrafish as a model organism. My introduction will begin with a description of the development of the vertebrate nervous system, with the main focus being on the developing zebrafish embryo. Subsequently, I will describe the features of Plzf and review details of its expression and function within various tissues. Finally, I will end with a detailed description of our current understanding of the role of Plzf within the CNS and the gaps in our knowledge that my PhD project aims to address.

1.1 Vertebrate Neurogenesis

The vertebrate nervous system is hugely complex and a complete understanding of its development is extremely challenging. It is comprised of hundreds of cell types that must be produced at the right time and position and organised into functional neuronal circuits. Throughout development a pool of cells must remain available in order to generate the required number and variety of cells during neuro- and gliogenesis. Our knowledge of the mechanisms behind the tight regulation of neurodevelopment has increased in recent years and remains a topic under intense investigation. In this section, I aim to give an overview of the key mechanisms involved in generating the required diversity of the CNS.

Formation of the nervous system

The complexity of the nervous system arises progressively and beings with a uniform field of epithelial cells gaining competence to become neurons or glia, via a process that is termed neural induction. At the onset of gastrulation, signals from the mesoderm are secreted to induce or inhibit neural induction within the ectodermal layer (Lumsden and Krumlauf, 1996). Early experiments in amphibians led to the development of the 'default model' for neural induction, which proposes that ectoderm is pre-programmed towards a neural fate unless inhibited by bone morphogenetic proteins (BMPs) (Vasquez et al., 2001). Important advances in our

understanding of neural induction have revealed that the process is considerably more complex and involves wingless-integrated (Wnt), fibroblast growth factor (FGF) and insulin-like growth factor family members (Wilson et al., 2002, Bally-Cuif and Hammerschmidt, 2003). Concomitant with neural induction is the expression of the SoxB family of transcription factors, which have been shown to be key regulators of neural specification (Pevny and Placzek, 2005).

These competent ectodermal epithelial cells form the neural plate and go on to be transformed into the neural tube in a process known as neurulation. The zebrafish neural plate can first be observed during early somitogenesis as a prominent thickening of cells (Kimmel et al., 1995). Early studies suggested that the neural plate is a monolayer (Papan and Campos-Ortega, 1994), whilst more recent work has shown that in some parts of the embryo the neural plate can be up to six cells deep (Tawk et al., 2007, Clarke, 2009). Neurulation ends with the formation of the neural tube, however the steps involved vary between different organisms.

The majority of vertebrates undergo a mechanism whereby the lateral edges of the neural plate move towards each other and fuse at future dorsal midline, forming the hollow neural tube (Blader and Strähle, 2000). In zebrafish, however, the plate initially becomes a solid structure termed the neural keel, which subsequently rounds into a cylindrical solid neural rod before finally becoming the neural tube through cavitation (Papan and Campos-Ortega, 1994). This process is shown schematically in Figure 1-1.

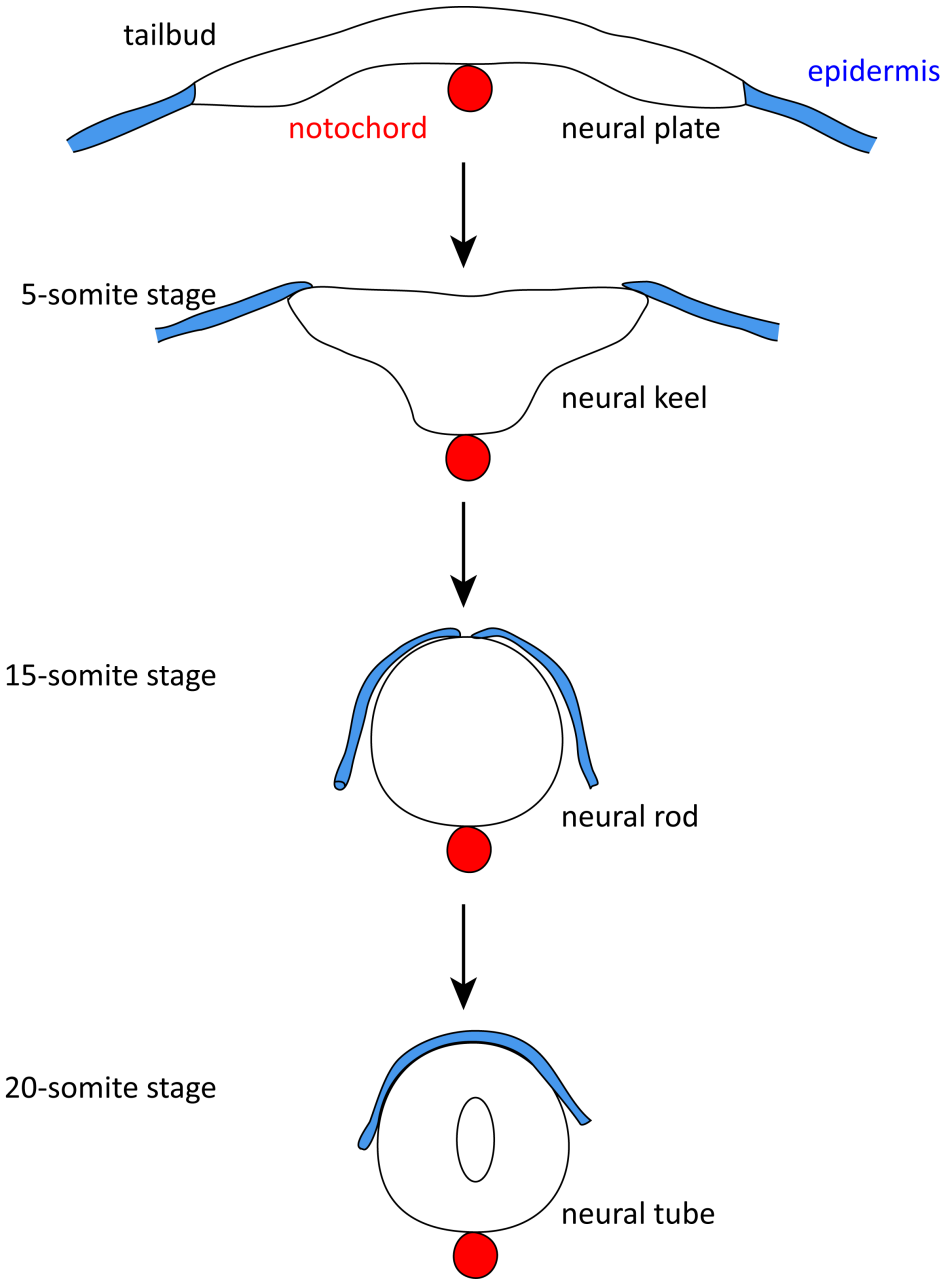
In all vertebrates, the neural tube is patterned along the three body axes: the anterior-posterior, dorsal-ventral and left-right axes; providing the progenitor cells with positional information that gives the cell its correct identity (Altmann and Brivanlou, 2001). Following neural induction, cells of the neural plate display an anterior fate and are transformed to adopt a posterior character by various signalling molecules (Wilson et al., 2002). Several discrete local organisers act to divide the neural plate into four anatomical divisions along the anterior-posterior axis, providing regions with distinct spatial identity (Kiecker and Lumsden, 2012). At the anterior, the brain is made up of the forebrain, midbrain and hindbrain while the posterior neural tube forms the spinal cord. The hindbrain is further subdivided into seven or eight evolutionarily conserved, lineage-restricted compartments, known as

Figure 1-1: Formation of the neural tube

Cartoon depicting the major steps in the formation of the neural tube in zebrafish. Images shown are cross-sections, with dorsal to the top. A thin layer of neuroepithelial cells make up the neural plate, observed at tailbud stage. The notochord is positioned medially and ventral to the neural plate, shown in red. Neural plate cells converge towards the dorsal midline, forming initially the neural keel and later the neural rod. By the 20-somite stage, the neural rod has formed a cavity and the neural tube is formed.

Figure adapted from Papan and Campos-Ortega (1994).

Figure 1-1


tailbud
epidermis
notochord
neural plate
5-somite stage
neural keel
15-somite stage
neural rod
20-somite stage
neural tube

rhombomeres (Lumsden and Krumlauf, 1996). Rhombomere identity is conferred by the restricted expression of specific transcription factors such as the *Hox* genes. Providing rhombomeres with individual identity is required to ensure that the correct pattern of neural subtypes arise later in development (Kimmel, 1993, Pasini and Wilkinson, 2002, Kiecker and Lumsden, 2005).

The neural tube is also patterned along its dorsal-ventral axis. The most-ventral cell type in the neural tube is the floor plate, whilst cells in the dorsal neural tube initially give rise to neural crest cells and subsequently to roof plate cells (Tanabe and Jessell, 1996). The generation of ventral cell types comes from inductive signals originating from the notochord, a mesodermal structure positioned medially and ventrally to the neural tube. A major inductive signal is Sonic hedgehog, initially acting to establish the floor plate as a source of the morphogen and subsequently specifying functionally distinct neurons at discrete positions along the dorsal-ventral axis (Blader and Strähle, 2000, Dessaud et al., 2008). BMP signalling molecules from the epidermal ectoderm induce the formation of the roof plate (Chizhikov and Millen, 2004), which later produces Wnt and BMP signalling molecules essential for the patterning of dorsal cell types (Chizhikov and Millen, 2005).

Generating the cells of the nervous system

Cells of the neural plate divide symmetrically, expanding the pool of progenitors required to generate the correct number of differentiated neural cell types. These bipolar neuroepithelial cells span the width of the neural plate, undergoing interkinetic nuclear migration whereby the cell nucleus moves between the basal and apical surface synchronised with the cell cycle (Takahashi et al., 1993). Cells undergoing mitosis are positioned at the apical surface whilst DNA synthesis occurs when the cells are located in more basal positions.

During neural tube formation, the multipotent neuroepithelial cells become radial glial cells, expressing astroglial markers such as glial fibrillary acidic protein (GFAP). These bipolar cells are characterised by their distinct morphology, their cell bodies remaining adjacent to the lumen of the neural tube and extending long glial-like fibres to make contact with both the apical surface and the membrane

surrounding the neural tube, termed the pial surface (Rakic, 1972). Similar to the neuroepithelial cells, radial glial cells undergo interkinetic nuclear migration, however this migration is restricted to an apical region referred to as the ventricular zone (Malatesta et al., 2008).

For many years, these cells were believed to have a purely structural role, existing as scaffolds to help guide migrating neurons to occupy their final position within the neural tube. At the turn of this century, it was determined that these cells are mitotically active and function as neural progenitors (Noctor et al., 2002). The prevailing view is that radial glial cells are the main source of neurons and glia within the CNS of vertebrates (Alvarez-Buylla et al., 2001, Malatesta et al., 2008). In mammals, the radial glial cells disappear shortly after birth, giving rise to ependymal cells and astrocytes, some of which continue to act as neural stem cells (Doetsch, 2003). In other vertebrates, such as the zebrafish, radial glial cells persist into adulthood and continue to generate new neurons (Rothenaigner et al., 2011).

The requirement to build a huge variety of neural cell types throughout embryogenesis necessitates that progenitor cells differentiate into neurons whilst maintaining a population of undifferentiated stem cells. Upon cell division, the progeny can either withdraw from the cell cycle and become irreversibly committed to a neuronal fate or remain as pluripotent, undifferentiated neural progenitors with the ability to self-proliferate (Bertrand et al., 2002). The production of neurons can be accomplished by asymmetric divisions that generate a neuron and another progenitor cell or by divisions that generate two differentiated neurons; both processes have been shown to be important in vertebrates and invertebrates (Roegiers and Jan, 2004). Upon division, the cell destined to become a neuron migrates basally along the fibres of the radial glial cells, until reaching its final position in the neural tube outside of the ventricular zone, an area termed the mantle zone. A simplified cartoon outlining the production of cells of the nervous system within the zebrafish is shown in Figure 1-2.

Asymmetric cell divisions of the neural stem cells have been extensively studied in invertebrates. Most neural stem cells, termed neuroblasts, divide asymmetrically to produce another self-renewing neuroblast and a ganglion mother cell (Brand and Livesey, 2011). The ganglion mother cell will divide once more,

producing two post-mitotic neurons or glial cells of a related identity that is dependent upon their birth order (Lin and Lee, 2012). Polarity proteins within the neuroblast function to promote asymmetric distribution of cell fate determinants, such as Prospero, Brat and Numb, into the differentiating daughter cell, ensuring that the apical daughter retains its proliferative capability (Wodarz and Huttner, 2003).

Whilst our understanding of these processes in vertebrates is less complete, studies have shown that radial glial cells are predominantly undergoing asymmetric cell divisions (Noctor et al., 2004, Alexandre et al., 2010) and that homologues of both the polarity proteins (Bultje et al., 2009, Alexandre et al., 2010) and cell fate determinants (Schwamborn et al., 2009) are involved. In the mouse neocortex, the daughter cell that inherits the greater amount of the polarity protein Par3 develops high Notch activity and remains a progenitor; whilst the other cell goes on to differentiate (Bultje et al., 2009). There is evidence that this mechanism is not universally true; studies in zebrafish have shown that Par3 segregates asymmetrically into the apical daughter along with the neurogenic protein Mind bomb (Alexandre et al., 2010, Dong et al., 2012). It is this apical daughter cell that goes onto differentiate whilst the more basally located daughter adopts the self-renewal fate (Figure 1-2).

In recent years, studies within both rodents and primates have revealed the existence of further classes of progenitors not found within lower vertebrates such as the zebrafish, suggesting that there is an evolutionary requirement to expand the production of neural cell types as brain size increases (Fish et al., 2008, Florio and Huttner, 2014). These include classes of radial glial cells that divide within the ventricular zone, but away from the apical surface (Pilz et al., 2013). Intermediate progenitors constitute a distinct progenitor cell type that arise from radial glial cell divisions at the apical surface and accumulate in a region known as the subventricular zone within the neocortex (Noctor et al., 2004). These cells are mitotically active and capable of undergoing one further round of symmetrical proliferative division away from the apical surface before being self-consumed by a symmetrical neurogenic division.

Figure 1-2: Neurogenesis in the developing vertebrate CNS

Simplified overview of neurogenesis in the embryonic CNS. Early neural progenitor cells are polarised neuroepithelial cells that undergo interkinetic nuclear migration in time with the cell cycle. When the nucleus is at the apical surface, the cells divide symmetrically, expanding the pool of progenitors. These neuroepithelial cells transition into radial glial cells which contain their cell bodies within the ventricular zone (orange) and extend glial processes through the mantle zone (white), contacting the basal (pial) surface. Radial glial cells can undergo asymmetric divisions, whereby one daughter cell becomes a neuron and the other will remain a progenitor. In the zebrafish, asymmetric inheritance of polarity proteins and Mind bomb into the apical daughter (blue) bias the neurogenic fate. This committed cell will migrate along the glial fibre, undergoing the neurogenic cascade (see text) until terminal differentiation in the mantle zone. As well as asymmetric divisions, radial glial cells can undergo symmetrical self-consuming divisions, producing 2 neurons. Over time, the ventricular zone narrows and the mantle zone expands.
Figure adapted from Paridaen and Huttner (2014).

Figure 1-2

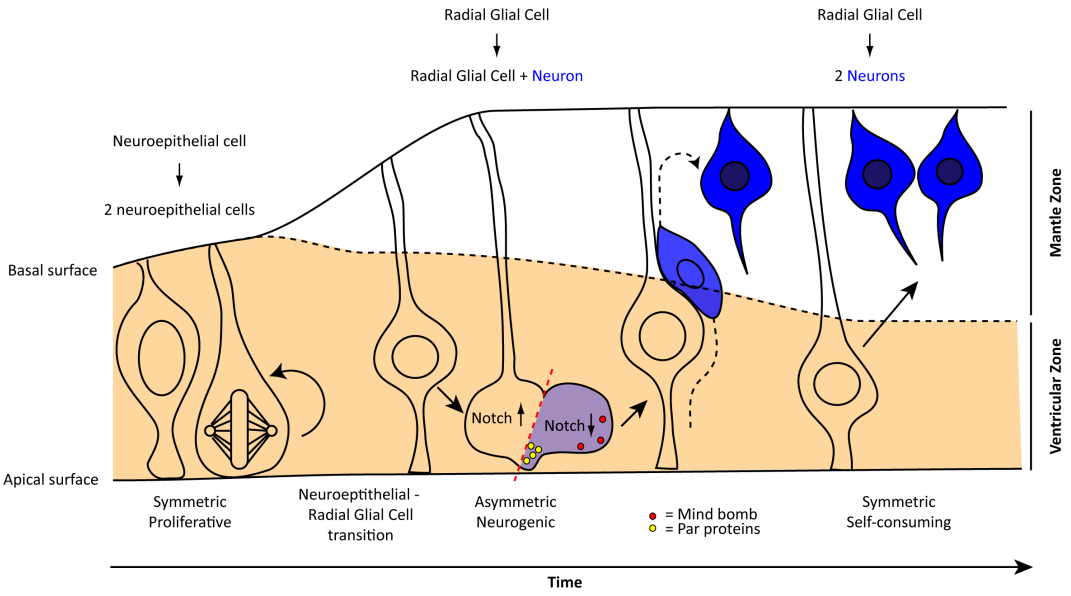
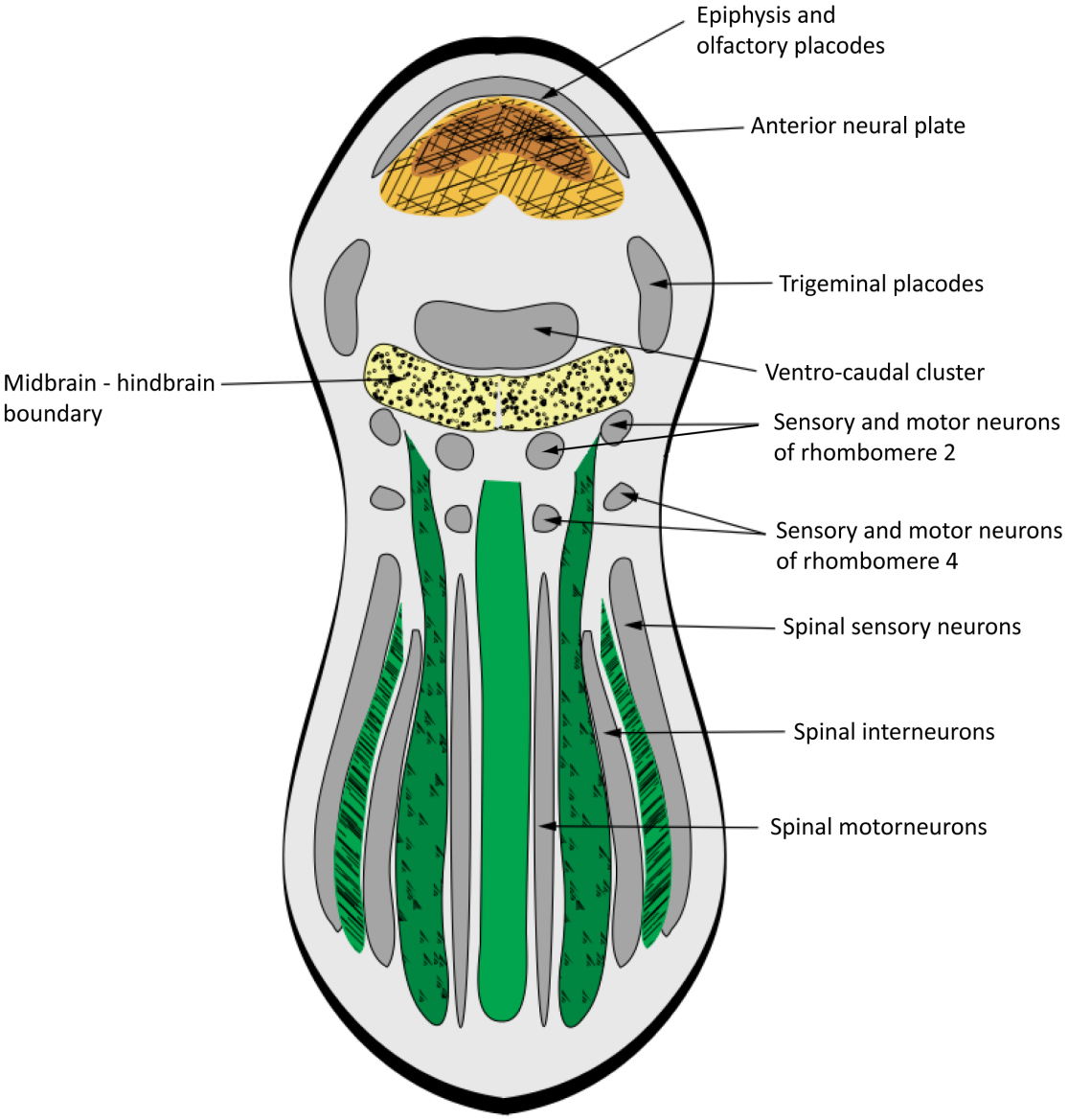

Radial Glial Cell
Radial Glial Cell + Neuron
Radial Glial Cell
2 Neurons
Neuroepithelial cell
2 neuroepithelial cells
Basal surface
Apical surface
Mantle Zone
Ventricular Zone
Notch
Notch
Symmetric Proliferative
Neuroeptithelial - Radial Glial Cell transition
Asymmetric Neurogenic
= Mind bomb
= Par proteins
Symmetric Self-consuming
Time

Figure 1-3: Patterning of neurogenesis

Spatial patterning of neurogenesis at the early neural plate stage in the zebrafish. A dorsal view of the embryo is shown, with anterior to the top. Labelled domains in grey are actively undergoing neurogenesis. Coloured patterns distinguish regions of the neural plate where neurogenesis is being actively inhibited. Figure adapted from Bally-Cuif and Hammerschmidt (2003).

Figure 1-3


Epiphysis and olfactory placodes
Anterior neural plate
Trigeminal placodes
Ventro-caudal cluster
Midbrain - hindbrain boundary
Sensory and motor neurons of rhombomere 2
Sensory and motor neurons of rhombomere 4
Spinal sensory neurons
Spinal interneurons
Spinal motorneurons

Patterning of neurogenesis

Precursors within the neural plate will differentiate into neurons or glia in a temporally and spatially controlled manner. Several transcription factors that are downstream of the inductive signals have been identified, which specify broad regions of the neural plate as competent to adopt a neural fate. These include members of the Sox, Pou, Iroquois and Gli family of proteins (Sasai, 1998). Most of these factors are believed to act as transcriptional activators, such as Iro1 and Iro7, which promote expression of proneural genes in the anterior zebrafish brain (Itoh et al., 2002).

After neural plate induction, large-scale spatial patterning events generate discrete neurogenic and non-neurogenic zones in the embryo (Figure 1-3). This is controlled by several spatially restricted inhibitory factors that prevent cells from embarking on a neurogenic fate. The anterior neural plate and midbrain-hindbrain boundary are two well-studied regions in which neurogenesis is inhibited in all vertebrates (Bally-Cuif and Hammerschmidt, 2003). These cells are retained as slowly dividing progenitors and are characterised by expression of various transcription factors such as Notch-independent members of the Hairy / Enhancer of Split (Hes/Her) families (Geling et al., 2003), members of the *zic* gene family (Brewster et al., 1998), or forkhead domain transcription factors (Bourguignon et al., 1998, Hardcastle and Papalopulu, 2000), which function to define domains of neurogenesis (Stigloher et al., 2008).

bHLH transcription factors involved in promoting neurogenesis

Both the specification of competent cells and initiation of neurogenesis rely on a group of basic helix-loop-helix (bHLH) proneural transcription factors (Bertrand et al., 2002). As a result of the pre-patterning of the neural plate, neurogenesis is restricted to cell clusters defined by the expression of proneural genes, indicated in grey in Figure 1-3. Two distinct families of proneural genes exist in vertebrates: those such as *ascl1*, which are related to the *Drosophila achaete-scute* genes; and those related to *Drosophila atonal* such as the *neurogenins* (*neurog*) (Brunet and Ghysen, 1999). These transcription factors specifically bind DNA sequences containing the E-box motif as heterodimeric complexes along with the ubiquitously

expressed E proteins in both vertebrates and invertebrates (Massari and Murre, 2000). In various organisms, ectopic expression of proneural genes is sufficient to cause cells to upregulate pan-neuronal gene expression and adopt a neuronal fate. (Ma et al., 1996, Farah et al., 2000, Mizuguchi et al., 2001).

A member of the *hes* family of bHLH genes, *hes6*, has a unique role in promoting differentiation. *hes6* expression is independent of Notch signalling, described below, and is expressed downstream of proneural proteins (Seo et al., 2007, Koyano-Nakagawa et al., 2000, Bae et al., 2000). Despite sharing structural similarity with other bHLH transcription factors, Hes6 is able to promote neuronal differentiation without binding DNA (Koyano-Nakagawa et al., 2000, Bae et al., 2000). Instead, this protein was found to inhibit the activity of the inhibitory Hes1 protein at a post-translational level. Hes6 therefore mediates a positive-feedback loop with the proneural genes to ensure the selected cells follow their path of differentiation.

Pathway to differentiation – the neurogenic cascade

In order to promote neurogenesis, the proneural genes must initially inhibit expression of the SoxB1 family of genes which, having been expressed as a result of neural induction, have functionally redundant roles in keeping cells undifferentiated (Pevny and Placzek, 2005). Proneural proteins have been shown to do this by upregulating expression of another Sox gene, *sox21*, which counteracts the activity of the SoxB1 family (Sandberg et al., 2005).

Expression of the proneural genes, such as the *neurogenins*, is transient and does not overlap with markers of terminally differentiated neurons (Ma et al., 1996, Gradwohl et al., 1996). The ability of these proneural genes to induce neural differentiation has been shown to involve subsequent events that require activation of downstream regulatory genes (Bertrand et al., 2002).

Many of the genes downstream of these early proneural genes are structurally- and functionally-related bHLH genes, which themselves are transiently expressed and partially overlap at successive stages of differentiation (Perron et al., 1998, Lee et al., 1995). The sequential expression is unidirectional (Ma et al., 1996) and is reminiscent of similar cascades occurring for bHLH proteins in muscle

development (Jan and Jan, 1993). In vertebrates, members of the NeuroD family of proteins function as part of this cascade and have the capacity to promote neuronal differentiation when ectopically expressed (Lee et al., 1995). *neurod4* (also known as *zath3* or *neurom*) and *neurod* are expressed sequentially in cells lining the ventricular zone that have exited the cell cycle and are beginning to migrate to their final position (Roztocil et al., 1997, Wang et al., 2003). These genes are known to be direct transcriptional targets of the Neurogenins (Seo et al., 2007) and can cross-activate each other but cannot activate expression of *neurogenin* (Perron et al., 1999). The onset of neuronal differentiation also results in the upregulation of the *delta* genes (Haddon et al., 1998).

Following expression of the proneural genes, postmitotic neurons express genes such as the RNA-binding proteins of the Hu family, which can be used as a pan neuronal marker (Kim et al., 1996a, Kim et al., 1997). These proteins promote neuronal maturation by operating on several downstream targets, acting both to increase the stability of mRNA molecules and to promote protein synthesis (Pascale et al., 2008).

bHLH transcription factors involved in inhibiting neurogenesis

Not all bHLH genes have a role in inducing neuronal differentiation. There exists a set of bHLH genes that function as inhibitory transcription factors, including *hes1, hes3* and *hes5* that are homologues of the *Drosophila hairy* and *Enhancer of split* genes and are able to repress neuronal differentiation (Kageyama et al., 2005). Members of this class of bHLH transcription factors bind to target genes such as *ascl1* along with the corepressor Gro/TLE, homologous to *Drosophila* Groucho, and repress their transcription (Grbavec and Stifani, 1996, Chen et al., 1997). A second method by which the Hes proteins prevent differentiation is by competing with the proneural proteins for the E proteins (Sasai et al., 1992). Hes proteins form a non-functional heterodimer with the E proteins and therefore prevent the proneural transcription factors from inducing neurogenesis.

Keeping cells undifferentiated – Notch-mediated lateral inhibition

In differentiating neurons, proneural factors drive the expression of the transmembrane ligands Delta or Serrate, which are able to bind and activate the

Notch receptor on adjacent cells through direct contact (Bray, 2006). Binding of the ligand results in the γ-secretase enzyme cleaving the Notch intracellular domain. Release of this intracellular domain causes it to translocate to the nucleus where it forms a complex with the suppressor of hairless protein, or its vertebrate homologue RBPj, altering its function from a transcriptional repressor to an activator of the inhibitory *hes* genes (Kageyama et al., 2005). This mechanism is known as lateral inhibition and allows selected cells to indirectly positively regulate their own proneural gene expression and prevent surrounding cells from adopting the same fate (Figure 1-4a).

The importance of Notch signalling in generating the correct number and variety of neurons has been demonstrated in numerous organisms. In the mouse CNS, inactivation of the 3 major Notch-dependent *Hes* genes: *Hes1, Hes3* and *Hes5,* causes premature neuronal differentiation resulting in the depletion of progenitor cells (Hatakeyama et al., 2004). They show that these proteins do not affect the initial formation of neural epithelial cells, but rather have an integral role in their maintenance. Abolishing lateral inhibition through the addition of a chemical blocker of Notch activation results in a significant increase in early-born neurons (Geling et al., 2002).

The classical understanding of lateral inhibition came from studies in the *Drosophila* nervous system where proneural genes are expressed in clusters of undifferentiated ectodermal cells (Artavanis-Tsakonas et al., 1999). The model assumes all cells are equivalent at first but stochastic differences in levels of proneural gene expression arise and result in a 'salt and pepper' pattern, where undifferentiated cells surround cells exhibiting high proneural gene expression. Lateral inhibition amplifies proneural expression in the committed cell until it reaches the threshold required to differentiate, at which point it delaminates from the surrounding ectoderm, relieving the inhibition on surrounding cells and a new cell is able to start the process of lateral inhibition again (Figure 1-4c).

A similar expression pattern was observed within the vertebrate nervous system, leading to the hypothesis that the same processes were involved (Chitnis et al., 1995). This model has recently been challenged by studies suggesting that proneural gene expression is more dynamic than previously realised. Real-time

imaging of neural progenitors has revealed that expression of the *hes1* transcript and protein oscillates in individual, proliferating cells (Shimojo et al., 2008). In these same cells, the proneural gene *neurog2* and Notch ligand *delta-like 1* oscillate with an inverse correlation to that of *hes1* (Shimojo et al., 2008, Imayoshi et al., 2013). These results argue that the 'salt and pepper' pattern observed is due to these dynamic changes in expression, rather than due to the amplification of stochastic differences. These oscillations are essential for maintaining cells in a proliferative state (Imayoshi et al., 2013) and once differentiated, Hes1 expression disappears and proneural gene expression becomes sustained. These oscillations were not observed at non-neurogenic boundary regions, where Hes1 expression was stable and at high levels (Shimojo et al., 2008).

Hes1 is able to repress its own transcription by directly binding to its promoter, which, combined with the relatively short half-life of the protein and mRNA, results in the oscillatory expression (Hirata et al., 2002). The microRNA miR-9 has been shown to play an important role in both the maintenance and termination of these oscillations (Bonev et al., 2012, Goodfellow et al., 2014). The primary miR-9 transcript (pri-miR-9) oscillates out of phase with Hes1 due to mutual repression of the two components, but the mature version of the microRNA is highly stable and accumulates over time. This presents a model whereby there is a gradual increase in mature miR-9 as the cells oscillate, eventually reaching a critical level that allows differentiation into neurons (Bonev et al., 2012) (Figure 1-4b). Furthermore, this model also provides a mechanism by which the progenitor cells can switch between different states, such as the slowly dividing progenitors in regions that include the midbrain-hindbrain boundary, and cells actively undergoing neurogenesis (Goodfellow et al., 2014). The revised model of lateral inhibition is shown in Figure 1-4d.

Several lines of evidence indicate that endocytosis of the Notch ligands is essential for correct receptor activation. Ubiquitin ligases such as Mind bomb have been identified as being key regulators of ligand signalling activity. The zebrafish mutant for Mind bomb displays a dramatic neurogenic phenotype where progenitors differentiate prematurely without giving rise to the full spectrum and number of neurons and glia (Itoh et al., 2003).

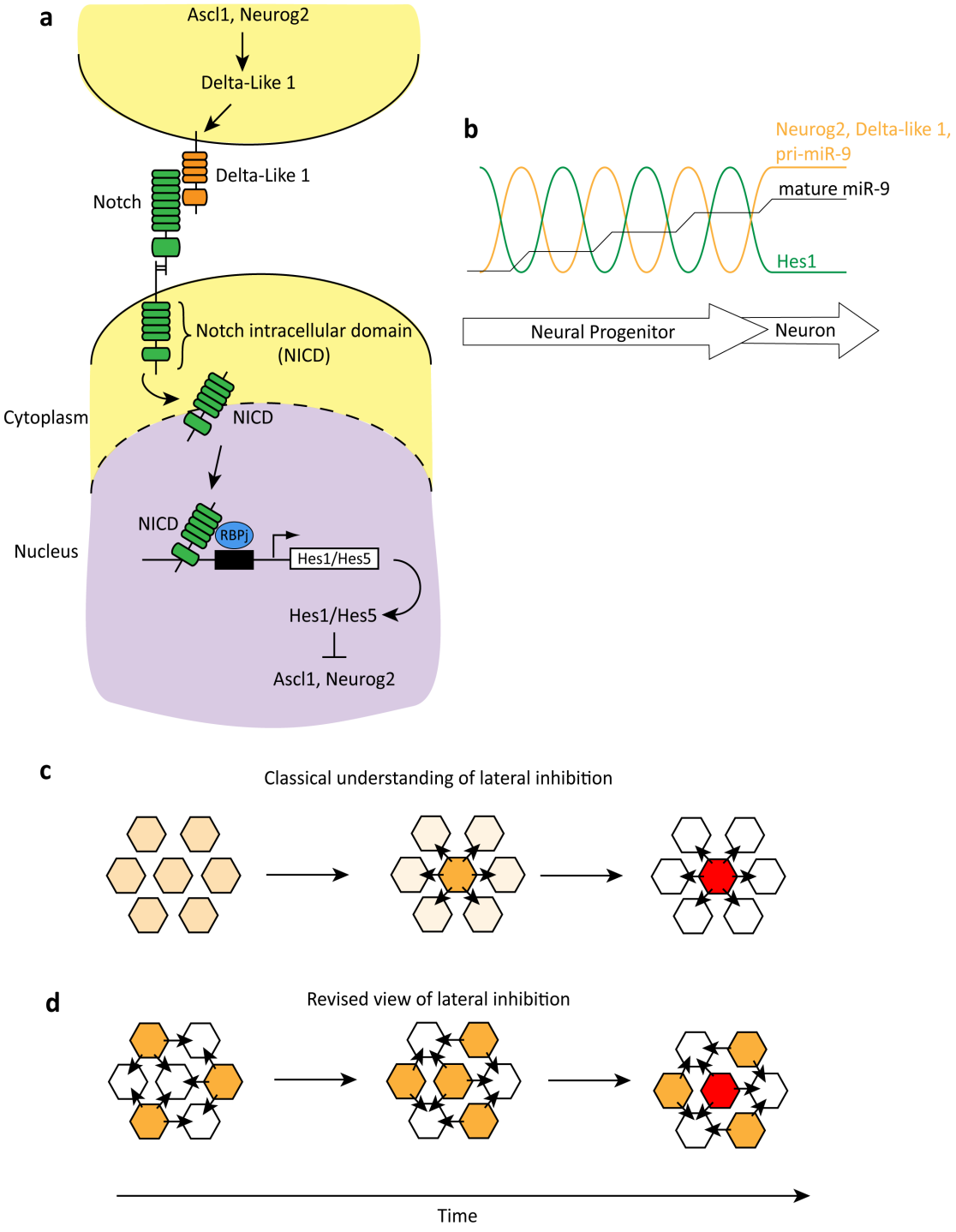
Figure 1-4: Notch signalling during neurogenesis

a: Overview of Notch signalling in vertebrates. Proneural factors Ascl1 and Neurog2 induce expression of the cell surface Notch ligand Delta-like 1. This cell interacts with and activates Notch, expressed on the surface of a neighbouring cell. Activation of Notch results in the release of the intracellular domain (NICD), which is transported to the nucleus where it forms a complex with RBP-J. This complex drives expression of the inhibitory *hes* genes, *hes1* and *hes3,* which prevent expression of proneural genes through transcriptional and post-translational mechanisms as described in the text. Figure adapted from Kageyama et al. (2008).

b: Oscillatory dynamics of Notch components within a single progenitor cell. Expression of Neurog2, Delta-like 1 and pri-miR9 oscillate out of phase with Hes1 due to repressive mechanisms between the various components. Mature miR-9 is not affected and accumulates over time. The cell differentiates upon reaching a threshold of mature miR-9 expression, at which point oscillations stop and expression of Hes1 remains low and Delta-like 1 and Neurog2 remain high. Figure adapted from Shimojo et al. (2008) and Bonev et al. (2012).

c & d: Cartoon showing classical (c) understanding of lateral inhibition and revised (d) view with the dynamic expression factored in. Classically, within a field of equipotent cells, stochastic differences in proneural expression result in one cell (orange) having higher levels than surrounding cells. Lateral inhibition amplifies this expression resulting in this cell delaminating from the epithelium and adopting a neuronal fate (red). In the revised view, the salt and pepper pattern of proneural gene expression changes dynamically. This means that expression of proneural gene at a given point in time does not necessarily mean that this cell will go onto differentiate. Figure adapted from Shimojo et al. (2008).

Figure 1-4


a
Ascl1, Neurog2
Delta-Like 1
Delta-Like 1
Notch
Notch intracellular domain
(NICD)
Cytoplasm
NICD
NICD
RBPj
Nucleus
Hes1/Hes5
Hes1/Hes5
Ascl1, Neurog2
b
Neurog2, Delta-like 1,
pri-miR-9
mature miR-9
Hes1
Neural Progenitor
Neuron
c
Classical understanding of lateral inhibition
d
Revised view of lateral inhibition
Time

Members of the Fringe family of proteins have an important role in glycosylation of the Notch receptor, which has been shown to alter the affinity of Notch binding to its ligands (Moloney et al., 2000). Studies within the zebrafish have shown that one of the fringe proteins, Lunatic fringe, is downstream of proneural gene expression and acts to maintain sensitivity to lateral inhibition (Nikolaou et al., 2009).

Balance between proliferation and differentiation

Cell cycle progression and cell differentiation are highly interdependent and careful coordination of the two processes is essential in order to generate a functional nervous system. The transition from proliferating neural progenitors to post-mitotic cells is accompanied by a lengthening of the cell cycle and involves crosstalk between proneural factors and cell cycle regulators (Bertrand et al., 2002). Studies have shown that artificially shortening the cell cycle acts to delay neurogenesis and increase the expansion of progenitor cells (Lange et al., 2009). During proliferation, cyclin-mediated phosphorylation of the proneural protein Neurog2 inhibits its ability to bind DNA, whilst cell cycle lengthening permits an accumulation of the unmodified protein resulting in activation of downstream targets and cell differentiation (Ali et al., 2011).

As previously described, the dynamics of interkinetic nuclear migration is known to correlate with cell cycle progression. Experiments in the zebrafish retina have revealed that if this migration is perturbed, then premature differentiation occurs the expense of later-born interneurons and glia (Del Bene et al., 2008). The Foxg1 (XBF-1) transcription factor involved in setting up the spatial pattern of the neural plate is involved in regulating proliferation, separate to its role in neural cell fate specification (Hardcastle and Papalopulu, 2000). The authors find that low levels of Foxg1 expression is able to induce expression of the cell cycle inhibitor $p27^{XIC1}$ and cause cell cycle arrest, whilst high levels of Foxg1 suppress expression of the inhibitor and lead to increased cell proliferation.

For some time it has been known that proneural genes are able to directly bind genes associated with the inhibition of cell growth and therefore promote cell differentiation (Pagliuca et al., 2000, Bertrand et al., 2002). Surprisingly, recent work

has revealed that proneural protein Ascl1 acts to directly drive the proliferation of neural progenitors as well as their subsequent cell cycle exit and differentiation (Castro et al., 2011).

Further control of proneural gene expression and function

The mechanisms described so far represent a core neurogenic program that functions to restrict neurogenesis to specific loci within the developing neural epithelium and goes some way to ensuring the correct number and pattern of neural cells exist within the CNS. Several other mechanisms act to further refine neurogenesis by controlling both proneural gene expression and function.

Histone deacetylases (HDACs), described in more detail when discussing Plzf function later, are associated with altering the epigenetic landscape through the removal of acetyl groups from histones (Gallinari et al., 2007). Within the CNS, HDACs are expressed within the neural progenitor cells and function to promote the expression of proneural genes (Cunliffe, 2004). In animal models where HDAC function has been attenuated, neurons fail to differentiate and increased progenitor proliferation is observed (Cunliffe, 2004, Montgomery et al., 2009). Although HDAC activity is commonly associated with transcriptional repression, one way in which HDAC1 is able to induce neurogenesis in the zebrafish is by binding to the promoter of the proneural gene *ascl1b* and drive its transcription (Harrison et al., 2011).

The *id* family of genes encodes transcription factors with a HLH domain but which lack the basic domain required for DNA binding. They are able to compete with proneural proteins for the required E proteins, acting as widespread dominant-negative regulators (Ruzinova and Benezra, 2003). Attenuating Id protein activity leads to cell cycle exit and premature differentiation (Lyden et al., 1999).

As well as having a role in the patterning of the neural tube, secreted signalling molecules such as retinoic acid, FGF, Wnt and BMP family members, have been implicated in the promotion and inhibition of neurogenesis (Hardcastle et al., 2000, Maden, 2007, Ford-Perriss et al., 2001, Michaelidis and Lie, 2008, Bond et al., 2012). The involvement of these various molecules at different stages of neurogenesis is complex and the intricate details are still being researched, with studies demonstrating opposing roles for various family members (Borello et al.,

2008) and both cooperative and antagonistic actions between different molecules (Diez Del Corral et al., 2003). Both the signalling molecules and their receptors can be broadly expressed throughout neural and surrounding tissues; whilst in other cases they have a more restricted expression pattern. An example of the latter is the expression of *fgf20a* within a subset of neurons in the zebrafish hindbrain that activates the FGF receptor(s) within segment centres and is responsible for the inhibition of neuronal differentiation (Gonzalez-Quevedo et al., 2010).

The majority of the inhibitory mechanisms described so far act on the initial onset of neuronal differentiation, inhibiting proneural gene expression before the neurogenic cascade can begin. There is evidence for regulation at multiple steps of the cascade. For example, in addition to having the earlier described role in progenitor maintenance, the microRNA miR-9 acts to directly inhibit HuC expression in order to regulate timing of neuronal differentiation (Coolen et al., 2012). In the chicken spinal cord, two *Hes6* genes exist: *Hes6-1* and *Hes6-2* (Vilas-Boas and Henrique, 2010). *Hes6-2* expression overlaps with *neurog1* and acts to repress transcription of *Hes5*. *Hes6-1* is expressed in a successive stage of neurogenesis, overlapping with *Neurod4* and having a post-translational role in inhibiting Hes proteins similar to that previously described of for its murine homologue.

Generating variety in the CNS

Whilst the processes described so far go some way to explain how the correct number of neural cells are produced, they do not explain how the required variety of different subtypes are generated. Neural cells acquire distinct characteristics and fates provided by positional information along the body axes (Guillemot, 2007). This process is best understood within the vertebrate spinal cord, where neuronal identity is defined according to position within the dorsal-ventral axis (Jessell, 2000). Various bHLH proteins have been associated with providing neuronal identity in the developing CNS, frequently through the interactions with other neuronal-fate determinants (Bertrand et al., 2002).

As well as spatial patterning having a role in determining neural identity, temporal patterning is also involved. In vertebrates, different types of neurons are born in a stereotypical order in certain regions of the CNS (Jacob et al., 2008). In

Drosophila, sequential transcription factor expression in neuroblasts produces differentially fated ganglion mother cells, providing neurons with a unique identity (Isshiki et al., 2001). It seems likely that a similar process is occurring in vertebrates and details are beginning to emerge from studies in the mammalian cortex (Kohwi and Doe, 2013).

The best understood switch in cell fate known to occur in multiple regions of the vertebrate CNS is the transition from neurogenesis to gliogenesis. Neurons are produced before glia during development and this temporal control can be recapitulated *in vitro* (Qian et al., 2000). Several intrinsic and extrinsic signalling pathways have been implicated in this switch, including evidence that proneural genes inhibit pathways that activate glial-specific genes (Sun et al., 2001, Miller and Gauthier, 2007). Furthermore, evidence exists to suggest that Notch signalling itself promotes the differentiation of glial cell types at the expense of neuronal cell types (Tanigaki et al., 2001).

Neurogenesis in zebrafish

Neurogenesis in zebrafish, as with other lower vertebrates, has been classically separated into two waves that occur successively, classed as primary and secondary neurogenesis (Kimmel, 1993). Primary neurogenesis is initiated at late gastrulation and produces neurons with large cell bodies that extend long axons to form the first functional neuronal circuits within a day of embryogenesis (Kimmel and Westerfield, 1990, Grunwald et al., 1988). Examples of primary neurons produced include the Mauthner neurons, born in the centre of rhombomere 4 in the developing hindbrain, sensory neurons in the head and motor neurons in the spinal cord (Mendelson, 1986). These neurons form part of the simple neural circuitry that drives the startle response, initiated by a single Mauthner cell action potential in response to excitatory input from the sensory neurons, and resulting in activation of motor neurons and contraction of muscles surrounding the spinal cord (Kimmel and Westerfield, 1990). This circuit provides the early larva with the ability to respond to touch, therefore increasing its chances of survival and allowing subsequent neurogenesis to proceed.

Radial glial cells have been shown to arise at around 11 hours post fertilisation (hpf) in zebrafish, just as the initial neurons have differentiated (Kim et al., 2008, Marcus and Easter, 1995). As development continues, some of these primary neurons will be replaced during the wave of secondary neurogenesis (Wullimann, 2009). As an example, the group of mechanosensory neurons within the dorsal spinal cord classed as Rohon-Beard neurons degenerate gradually during larval stages and are replaced by the dorsal root ganglion neurons (Reyes et al., 2004). It should be noted that despite the classical definition of two separate waves, neurogenesis is a continuous process that is not obviously divided into two distinct phases (Lyons et al., 2003) and similar regulatory mechanisms are likely to play a role during both waves.

Studies have shown that neurogenesis within the adult mammalian brain is very limited and largely restricted to two regions, the subgranular zone in the dentate gyrus of the hippocampus and the subventricular zone of the lateral ventricles (Ming and Song, 2011). The adult zebrafish brain, however, is known to have much greater proliferative potential and continue to produce neurons throughout its life (Grandel et al., 2006). Radial glial cells persist in the adult, retaining their proliferative capability and ability to generate neurons, making them the most likely candidates for adult neural stem cells (Adolf et al., 2006, Pellegrini et al., 2007, Lam et al., 2009). Reflective of their continuous ability to generate neurons, zebrafish show a remarkable ability to regrow both axons and entire tissues following injury within the CNS (Becker and Becker, 2008).

The zebrafish hindbrain as a model for studying neurogenesis

In the hindbrain, segmental gene expression of Eph receptors and Ephrin ligands in complementary rhombomeres is essential to ensure cells from an individual rhombomere are unable to migrate into adjacent compartments (Cooke et al., 2005, Xu et al., 1995). Specialised boundary cells are observed to form at the interface of two rhombomeres and Eph-ephrin signalling is important for both the formation and maintenance of these cells (Cooke et al., 2005, Terriente et al., 2012).

The first hindbrain neurons to differentiate arise bilaterally in each rhombomere centre and subsequent neurons are added to the segments in a highly

stereotypical manner (Hanneman et al., 1988). By 24 hpf, clusters of neurons are observed within the mantle zone of the rhombomere centres and also close to the boundaries. Radial glial fibres are known to occupy the space adjacent to the rhombomere boundaries, where neurons are absent (Trevarrow et al., 1990). The cell bodies of these mitotically active glial cells are present at the ventricular surface while their processes are organised into two transverse rows, forming a structure known as the 'glial curtain' (Trevarrow et al., 1990, Marcus and Easter, 1995). This results in a highly stereotypical pattern of neurons and fibres within the mantle zone that is known to exist from 22 hpf and is maintained throughout embryonic development (Trevarrow et al., 1990).

Within the neural epithelium, cells actively undergoing neurogenesis are confined to zones flanking rhombomere boundaries within the hindbrain (Cheng et al., 2004). As previously described, signals from the *fgf20a*-expressing neurons act to prevent neurogenesis from occurring in rhombomere centres (Gonzalez-Quevedo et al., 2010). There is evidence that sustained Notch activation occurs at hindbrain boundaries and therefore this correlates with an inhibition of neurogenesis at this location (Cheng et al., 2004, Qiu et al., 2009). Recent work has shown that these boundary cells act as a source of the chemorepellent semaphorin family of proteins to position the *fgf20a* neurons in the segment centres (Terriente et al., 2012). The model of the role of boundaries and *fgf20a* neurons in the positioning of neurogenesis is shown in Figure 1-5. These components lead to a striking pattern of both neurons and progenitors cells within the zebrafish hindbrain that make it an attractive model for studying neurogenesis.

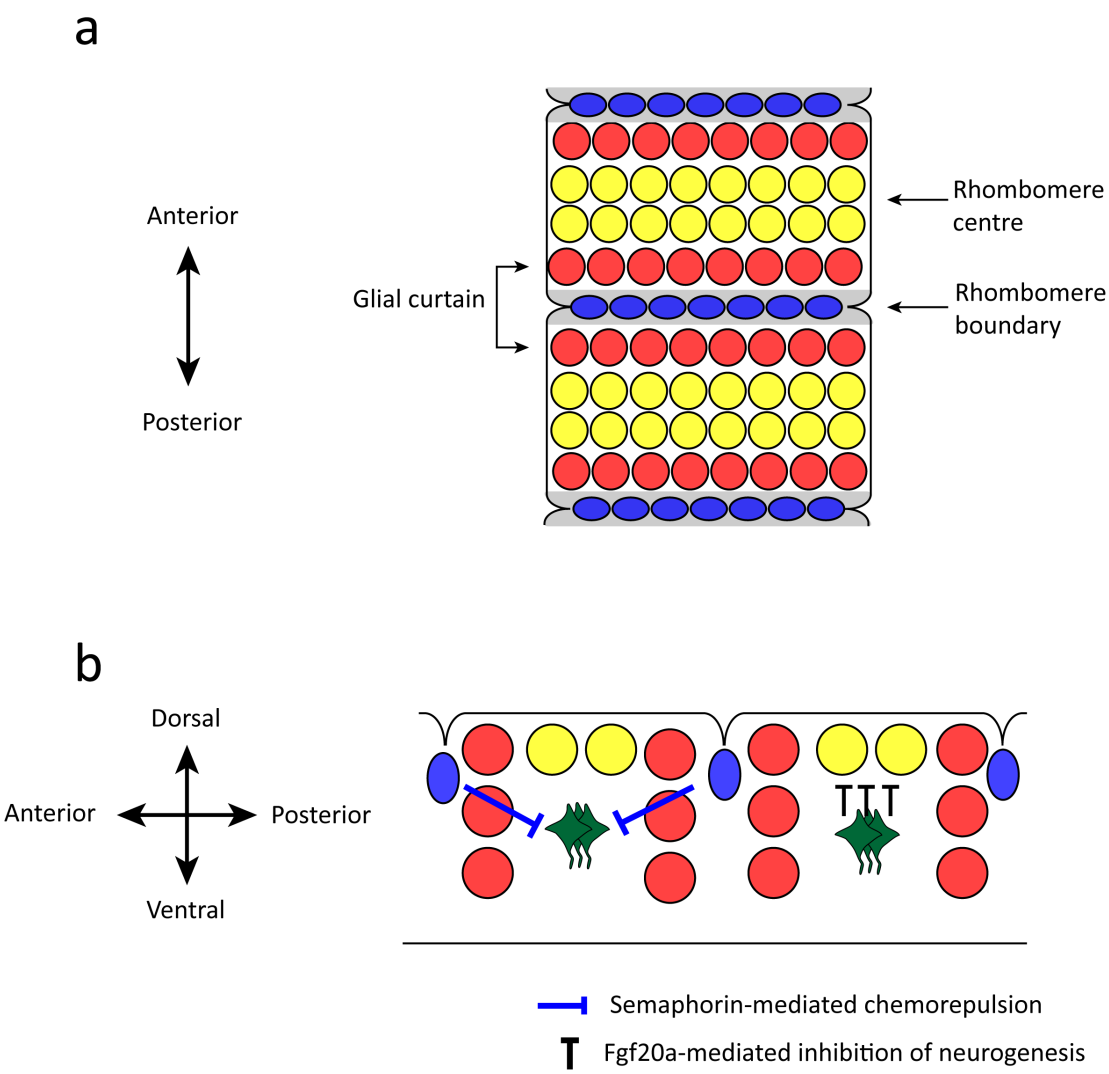
Figure 1-5: Neurogenesis in the zebrafish hindbrain

Simplified overview of neurogenesis in the zebrafish hindbrain.

a: Dorsal view of the zebrafish hindbrain showing the different classes of progenitor cells that exist at the ventricular surface. Cells undergoing neurogenesis (red) flank the rhombomere boundaries, forming a structure known as the glial curtain. Specialised boundary cells (blue) occupy positions at the rhombomere boundary, whist other cells (yellow) are maintained as undifferentiated progenitors in the rhombomere centres.

b: Sagittal (side) view of the zebrafish hindbrain. The boundary cells secrete the chemorepellent semaphorin family of proteins which position the *fgf20* neurons in the mantle zone of the segment centres. Signals from these cells act to prevent neurogenesis from being initiated in the cells that occupy the rhombomere centres.

Figure 1-5


a
Anterior
Posterior
Glial curtain
Rhombomere centre
Rhombomere boundary
b
Dorsal
Anterior
Posterior
Ventral
Semaphorin-mediated chemorepulsion
Fgf20a-mediated inhibition of neurogenesis
KEY
boundary cell
progenitor cell
cell undergoing neurogenesis
Fgf20a-neuron

1.2 The Promyelocytic Leukemia Zinc Finger Protein

The Promyelocytic Leukemia Zinc Finger (Plzf/Zbtb16) protein is a member of the zinc finger and BTB domain (Broad complex, Tramtrack and Bric-à-brac) containing family. Proteins of this family are characterised by the presence of an N-terminal BTB domain, also known as a POZ (Poxvirus and Zinc finger) domain, and several C-terminal Krüppel-type C_2H_2 zinc finger motifs. The human genome encodes 47 proteins of this family, which are generally classed as transcription factors and carry out a broad range of functions, both during development and within the adult (Privé et al., 2005). Plzf is one of the best studied of this family of proteins and in the following section I aim to give an overview of its known biological processes.

Functional domains of Plzf

In humans, the *PLZF* gene is located on chromosome 11 and composed of 7 exons. The gene contains several splice variants, of which the transcripts have been detected in multiple tissues (Zhang et al., 1999). The most common transcript is found in a large number of tissues and encodes a 673 amino acid protein, whilst the other transcripts encode N-terminally truncated Plzf proteins and have a more restricted expression pattern. In the mouse, the gene is located on chromosome 9 and only a single transcript is reported (Suliman et al., 2012). An isoform of human Plzf lacking the BTB domain has been described to have an important function in colorectal cancer cell lines (Jones et al., 2013).

The N-terminal 118 amino acids constitute the highly conserved BTB domain, which has a role in mediating protein multimerisation and is associated with transcriptional repression (Bardwell and Treisman, 1994, Dong et al., 1996). The crystal structure of the BTB domain of human Plzf has been solved, where it was found to exist as a homodimer (Ahmad et al., 1998, Li et al., 1999). This homodimerisation results in the formation of an interaction groove necessary to mediate transcriptional repression through the recruitment of corepressors, as described later (Hong et al., 1997, Melnick et al., 2002). Plzf has been shown to be able to form heterodimers with other BTB domain containing proteins *in vitro*, namely Bcl6 and Fazf (Dhordain et al., 2000, Hoatlin et al., 1999), but a function of these dimers has yet to be described.

Downstream of the BTB domain is a second repressive domain known as RD2. This domain recruits the corepressor eight-twenty one (ETO), a putative transcription factor with no known specific DNA binding activity but with a high affinity for nuclear corepressors (Melnick et al., 2000c). Mutations within this domain have been shown to affect the transcriptional repressive activity of Plzf (Kang et al., 2003). Several lysine residues around the RD2 domain can be sumoylated and mutating these residues significantly reduces the ability of Plzf to repress transcription (Chao et al., 2007).

As well as being a repressor, Plzf has been shown to bind directly to the promoters of specific genes and activate transcription (Hobbs et al., 2010, Doulatov et al., 2009). The mechanism through which Plzf can activate gene expression has not yet been elucidated. Early studies of human Plzf found a region between the BTB domain and RD2 domain that is rich in acidic residues and is capable of activating transcription when expressed alone (Li et al., 1997). This weak activating activity was masked by the presence of either the BTB domain or the RD2 domain and therefore its activity is attenuated in the complete protein.

The 9 C-teminal zinc finger domains of Plzf facilitate sequence-specific DNA binding to its target genes (Li et al., 1997). A consensus sequence for Plzf binding has been described as GTAC(T/A)GTAC (Ball et al., 1999). Studies have demonstrated that the DNA binding ability of Plzf is within the C-terminal 5 zinc fingers (Sitterlin et al., 1997) and *in silico* models predict that zinc fingers 6-8 account for binding to the consensus sequence, with the final finger important for enhancing high-affinity binding (Guidez et al., 2005). Acetylation of specific lysine residues within the final zinc finger by histone acetyl transferase p300 was shown to improve DNA binding (Guidez et al., 2005). As well as playing a role in DNA binding, these zinc fingers have been shown to participate in interactions with other proteins (Tsuzuki and Enver, 2002, Martin et al., 2003, Lin et al., 2013, Rho et al., 2007).

Plzf fusion proteins are involved in acute promyelocytic leukaemia

The human *PLZF* gene was initially identified due to its involvement in a rare form of acute promyelocytic leukaemia (APL) (Chen et al., 1993). This cancer of the blood and bone marrow is characterized by an abnormal accumulation of immature

Figure 1-6: Plzf domains and epigenetic modifications

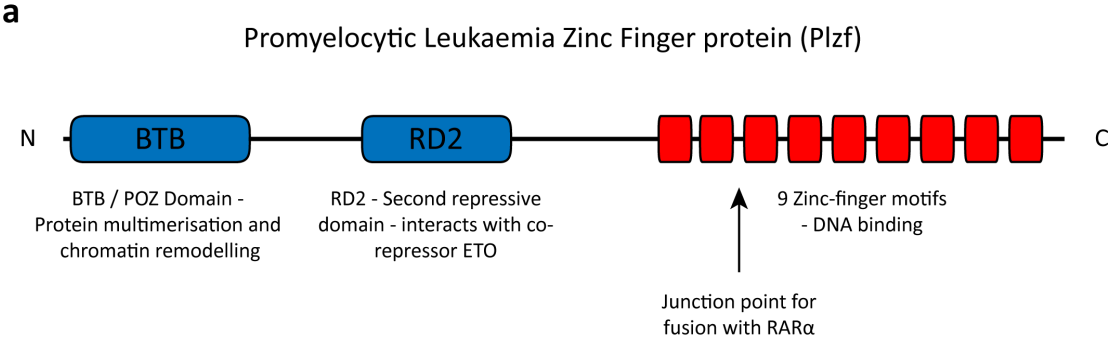
a: The Plzf protein is composed of several functional domains. At the N-terminus, the BTB domain is responsible for protein multimerisation and chromatin remodelling through the recruitment of corepressors. A second repressive domain, RD2, is located downstream of the BTB domain. At the C-terminus, 9 zinc finger motifs are responsible for sequence specific DNA binding. The arrow indicates the position in the protein that acts as a junction point for fusion with RARα in acute promyelocytic leukaemia.

b: Crystal structure of the BTB domain from two Plzf proteins forming a dimer. The arrow indicates the location of the charged pocket through which corepressors can bind. Figure adapted from Li et al. (1999).

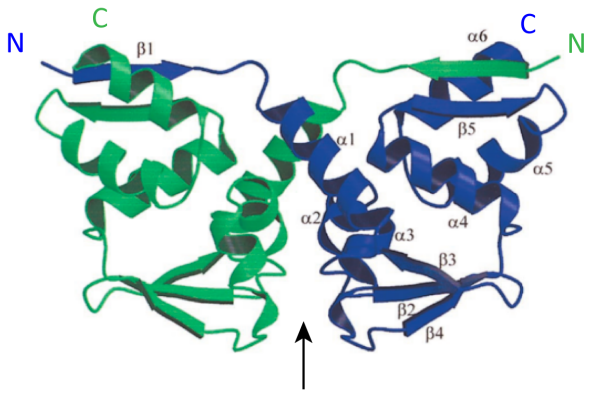
c: Cartoon depicting ability of Plzf to induce epigenetic modifications associated with transcriptional repression.

Figure 1-6

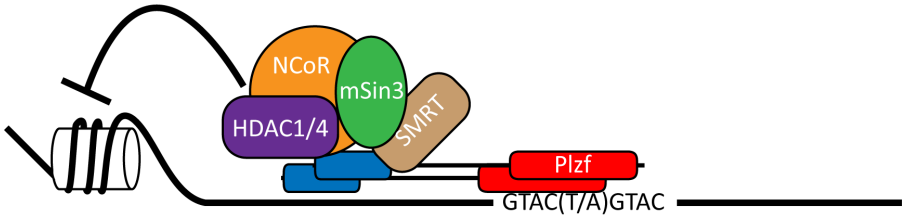
a



b



c



myeloid cells. It is caused by a chromosomal translocation that fuses part of the *retinoic acid receptor* α (*RAR*α) gene to one of five fusion partners, including Plzf (Lin et al., 2001) (Figure 1-6a). RARα acts as a transcriptional activator and the two reciprocal fusion products, Plzf-RARα and RARα-Plzf, repress genes normally activated by RARα and activate genes normally repressed by Plzf respectively (Lin et al., 2001).

Most forms of APL result in the recruitment of the repressive polycomb complex 2 to retinoic acid target genes (Villa et al., 2007). Patients with these forms of the disease are sensitive to treatment by all-*trans* retinoic acid, acting to release the complex from the target genes, and the patients undergo complete remission. Unlike typical APL, the transcriptional repression caused by the Plzf-RARα fusion product is resistant to this treatment (Licht et al., 1995). A potential explanation for this came from the finding that both Plzf and Plzf-RARα bind BMI-1 and recruit polycomb complex 1 to interact with retinoic acid response elements (Boukarabila et al., 2009). This stable, repressive complex is not removed following treatment with retinoic acid and patient prognosis is accordingly poor.

Plzf induces epigenetic modifications

Within the nucleus, DNA is packaged into nucleosomes by wrapping the negatively charged DNA strands around positively charged histone proteins. The folding of DNA within these structures can alter the accessibility to the genetic locus and therefore contribute to transcriptional control (Wolffe and Guschin, 2000). One of the major mechanisms by which this is affected is through changing the overall charge on the histones by dynamic addition and removal of acetyl moieties to lysine residues, carried out by two classes of enzymes: acetyl transferases and HDACs respectively. The action of HDACs results in tighter packing of DNA and is associated with transcriptional repression (Gallinari et al., 2007).

Early studies revealed that both Plzf and the fusion protein Plzf-RARα are able interact with HDACs and therefore it was hypothesised that it was through this mechanism that Plzf could repress transcription at target genes (Lin et al., 1998). The interaction groove formed upon Plzf dimerisation results in a highly conserved alignment of charged residues (Figure 1-6b), which are able to bind directly to

nuclear corepressors N-CoR and SMRT, which in turn can recruit HDAC1 and HDAC4 (David et al., 1998, Melnick et al., 2000a, Chauchereau et al., 2004) (Figure 1-6c). Furthermore, it has been demonstrated that interaction with ETO through the RD2 domain recruits HDAC proteins (Melnick et al., 2000b). The importance of forming these complexes to mediate transcriptional repression was demonstrated by showing that Plzf activity is sensitive to HDAC inhibitors (David et al., 1998) and by knocking down HDAC function (Chauchereau et al., 2004). In both studies it was found that some residual repressive activity remains, suggesting that Plzf is also able to repress transcription through other means.

Involvement of Plzf in the cell cycle

In many tissues, Plzf is associated with the inhibition of cell-cycle progression and differentiation, thus maintaining a quiescent state of progenitor cells. Cyclins are proteins that regulate the progression of cells through the cell cycle through the activation of cyclin-dependent kinases (Murray, 2004). Studies in cultured murine myeloid cells have demonstrated that Plzf represses cell cycle progression, causing cells to accumulate during S phase, and identified direct inhibition of the *CyclinA2* (*Ccna2*) gene as the cause (Shaknovich et al., 1998, Yeyati et al., 1999). Plzf is also able to specifically interact with retinoblastoma protein, leading to repression of both Plzf and retinoblastoma specific target genes involved in the cell cycle such as *Cdc6* (Petrie et al., 2008).

There is evidence to suggest that the role Plzf has in affecting proliferation is context and cell-type dependent. Despite the negative regulation of cell cycle progression by Plzf, studies have shown that Plzf is highly expressed in both quiescent and highly proliferative cells during hematopoietic development (Dai et al., 2002). Cyclin dependent kinase 2 (Cdk2) is able to phosphorylate Plzf at two separate sites, triggering its degradation and thereby promoting proliferation (Costoya et al., 2008), whereas Cdk1 has been shown to phosphorylate the same sites and bind to DNA in a complex with Plzf (Ball et al., 1999). Plzf has been shown *in vitro* to transiently bind the c-*Myc* promoter, where in different studies it has been shown to directly repress (Mcconnell et al., 2003) or drive expression (Doulatov et

al., 2009); however, *in vivo* experiments have failed to replicate these findings (Costoya et al., 2004).

Experiments in cultured myeloid cells have described an increase in apoptosis upon Plzf overexpression (Shaknovich et al., 1998). Further studies revealed a downregulation in anti-apoptotic genes coding proteins such as TERT, and an increase in inducers of apoptosis, such as Tp53inp1, Id1 and Id3 upon expression of Plzf (Bernardo et al., 2007). In cancer cell lines, overexpression of Plzf results in an upregulation of activated caspase-3, resulting in cell apoptosis (Rho et al., 2007). As well as this, experiments in lymphoid, spermatogonial and neural cells have demonstrated an increase in cell death upon loss of Plzf function (Costoya et al., 2004, Parrado et al., 2004, Gaber et al., 2013).

Animal models of Plzf function

A knockout mouse for *Plzf* was generated in 2000, in which the second exon was replaced with the neomyocin resistance gene (Barna et al., 2000). As well as this generated mutant, a mouse carrying the *luxoid* mutation that had spontaneously arisen 50 years previously was later mapped to the *Plzf* gene (Buaas et al., 2004). Both mice were viable and displayed patterning defects in both the limb and axial skeleton, including the presence of extra and transformed digits, and the males were found to be sterile (Barna et al., 2000, Buaas et al., 2004). A human patient with a biallelic loss of function of Plzf has been described (Fischer et al., 2008). The patient displayed clinical symptoms such as severe skeletal defects and genital hypoplasia reminiscent of the Plzf-deficient mice, along with mental retardation, craniofacial defects and microcephaly.

Plzf is expressed in a range of tissues

Expression of Plzf was first studied within the hematopoietic system. Haematopoietic stem cells arise within the bone marrow and generate cells of either the myeloid or lymphoid lineage (Kondo, 2010). Myeloid progenitor cells isolated from the bone marrow were found to express high levels of Plzf and expression declined as the cells became restricted towards terminal differentiation of most lineages (Reid et al., 1995, Doulatov et al., 2009). This was found not to be the case for megakaryocytic development, where Plzf expression increases during

differentiation, suggesting that Plzf can exert lineage-specific effects (Labbaye et al., 2002). Within the lymphoid lineage, Plzf expression is initially low until it is induced immediately after selection of the natural killer T cells, and subsequently decreasing during terminal differentiation (Savage et al., 2008, Kovalovsky et al., 2008).

During mouse embryogenesis, Plzf is expressed within the developing limb buds of both the forelimb and hindlimb (Barna et al., 2000). High levels of expression were found in the developing CNS, described in more detail later. During postnatal development, high levels of Plzf expression are observed within the quiescent spermatogonial progenitors; its expression later decreases during differentiation into mature sperm cells (Costoya et al., 2004, Hobbs et al., 2010).

Our knowledge of which genes are responsible for Plzf expression is limited. In developing natural killer T cells, the transcription factor Egr2 (Krox20) binds and activates the *Plzf* promoter (Seiler et al., 2012). Within this cell lineage, T cell receptor-mediated signalling has been implicated in inducing Plzf expression (Kovalovsky et al., 2008). Studies in human myeloid cells found that the zinc finger protein EVI-1 is essential for the activity of the *Plzf* promoter (Takahashi and Licht, 2002). Several conserved noncoding elements within the second intron of *Plzf* have been identified which, upon deletion within the rat genome, result in decreased expression in the limb buds without affecting the pattern of expression in the CNS (Liska et al., 2009).

Nuclear expression of Plzf

Plzf is expressed in cell nuclei, localised in a punctate pattern within nuclear bodies (Koken et al., 1997, Reid et al., 1995, Ruthardt et al., 1998). Consistent with a role in transcriptional regulation, removing Plzf from the nucleus results in a loss of its ability to repress target genes (Nanba et al., 2003, Doulatov et al., 2009, Sobieszczuk et al., 2010). Removing the BTB domain was shown to translocate Plzf from the nucleus and cause redistribution to the cytoplasm (Dong et al., 1996). Similarly, the isoform lacking the BTB domain is localised to the cytoplasm, where it has a role related to cell adhesion that is independent of transcriptional repression (Jones et al., 2013).

The addition of ubiquitin moieties to specific sites within the Plzf protein by posttranslational modification has also been shown to result in Plzf being equally distributed between the nucleus and cytoplasm (Kang et al., 2008). In myeloid cells, cytokines induced upon a stress response, such as interleukin 3, result in nuclear export and inactivation of Plzf (Doulatov et al., 2009). Plzf was found to physically interact with the zinc finger protein Sal-like protein 4 (Sall4) (Hobbs et al., 2012). When expressed alone, Sall4 and Plzf are localised to different regions of the chromatin within the nucleus. The authors found that if either protein is expressed at higher levels than the other, it was able to redirect the localisation of the corresponding protein. The expression level of either protein therefore dictates their localisation to different subnuclear domains.

Biological functions of Plzf

The striking patterning defects of the *Plzf* null mouse were accompanied by mis-expression of the *HoxD* gene complex in the hindlimb (Barna et al., 2000). Plzf was found to bind multiple sites within previously identified *cis* regulatory elements of *Hoxd11* (Barna et al., 2002). Plzf mediates long-range interactions by DNA looping, bringing the multiple binding sites spatially closer, and interacts with the Polycomb protein family member Bmi-1, resulting in chromatin remodelling (Barna et al., 2002, Boukarabila et al., 2009). Other members of the *HoxD* gene complex were shown to contain Plzf binding sites suggesting that Plzf controls the spatial expression of these genes in the limb bud.

In males, a pool of germline stem cells must remain available throughout sexual maturity in order to continue to produce sperm cells (Kotaja and Sassone-Corsi, 2004). Plzf was found to be restricted to the undifferentiated, progenitor-like spermatogonia of the testes that have arisen from the primordial germ cells during mouse embryogenesis (Costoya et al., 2004). The sterility of the knockout mouse was found to be due to a progressive loss of spermatogonia with age, caused by an initial burst of increased proliferative activity leading to a depletion of the germline stem cells (Buaas et al., 2004, Costoya et al., 2004). This suggests that Plzf is an essential regulator of spermatogonial stem cell maintenance and self-renewal (Kotaja and Sassone-Corsi, 2004).

Several targets of Plzf, both direct and indirect have been identified to be important during mouse spermatogenesis. The increase in proliferation upon loss of Plzf was linked to increases in expression of genes involved with the cell cycle, such as *Ccna2* and *Ches1* (Costoya et al., 2004). Plzf was found to directly bind and repress transcription of the stem cell factor receptor *kit* (Filipponi et al., 2007) and is able to drive transcription of *Redd1* (*Ddit4*), an inhibitor of the signalling complex mTORC1 (Hobbs et al., 2010). mTORC1 is a key mediator of cell growth, and aberrant activity of mTORC1 as a consequence of losing Plzf function results in exhaustion of the stem cell pool. The previously described interaction with Sall4 was found to have a functional role during spermatogenesis (Hobbs et al., 2012). Sall4 is associated with increased *kit* expression, having a mutually antagonistic role with Plzf that is essential for the maintenance of a stem cell pool.

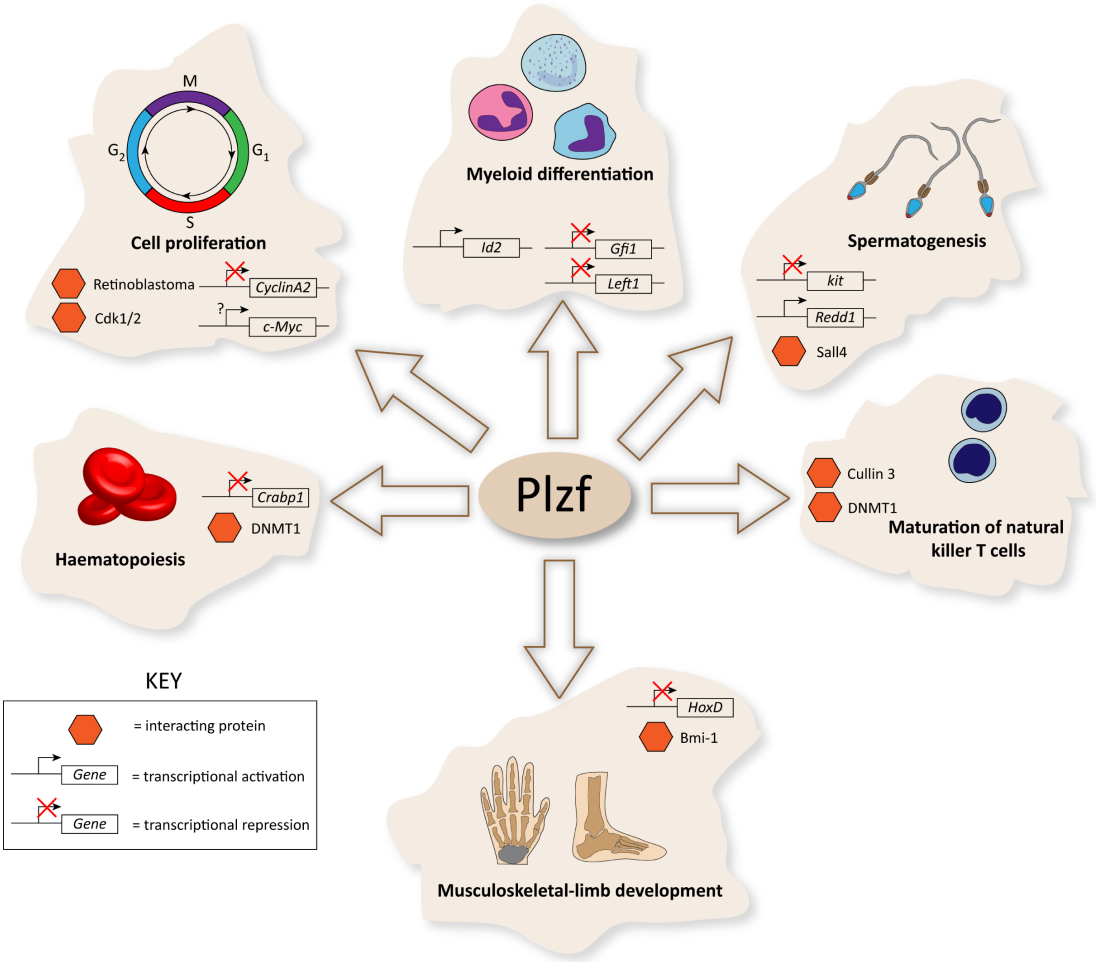
Forced expression of Plzf in myeloid cell lines results in an inhibition of proliferation and differentiation (Shaknovich et al., 1998). A more recent study makes use of a xenograft mouse transplanted with primary human myeloid cells, which had been transduced with Plzf overexpression or knockdown vectors (Doulatov et al., 2009). They showed that Plzf is essential to maintain the balance between progenitors and differentiated cells and that loss of Plzf results in a depletion of progenitors. It is able to do this by transcriptionally repressing genes associated with myeloid differentiation, such as *GFI1* and *LEF1*, and by activating transcription of inhibitors of differentiation such as *ID2*.

Experiments within human haematopoetic cells found that Plzf is able to directly bind to the *CRABP1* intron and cause transcriptional repression through deacetylation (Guidez et al., 2007). Alongside HDAC interactions, Plzf was found to also recruit the methyltransferase DNMT1 to the target site, resulting in a wave of methylation that advances towards to the promoter and causes gene silencing. Both HDAC and DNMT1 were found to interact with Plzf in the developing natural killer T cells, along with the ubiquitin ligase Cullin3 (Cul3) (Mathew et al., 2012). In this system, Plzf was found to transport Cul3 to the nucleus where it is able to add ubiquitin moieties to the chromatin-modifying complexes associated with Plzf.

Figure 1-7: Functions of Plzf

Plzf has been shown to be expressed and have roles in a number of different tissues. Shown here are some of its interacting partners and direct transcriptional targets that are known to be important for development of each indicated tissue. Not shown are the previously described corepressors and HDACs that have been demonstrated to cause transcriptional repression in each tissue.
Figure partially adapted from Suliman et al. (2012).

Figure 1-7


M
G2
G1
S
Cell proliferation
Retinoblastoma
Cdk1/2
CyclinA2
?
c-Myc
Myeloid differentiation
Id2
Gfi1
Lef1
Spermatogenesis
kit
Redd1
Sall4
Plzf
Haematopoiesis
Crabp1
DNMT1
Cullin 3
DNMT1
Maturation of natural killer T cells
KEY
= interacting protein
Gene
= transcriptional activation
Gene
= transcriptional repression
HoxD
Bmi-1
Musculoskeletal-limb development

1.3 Expression and function of Plzf within the CNS

Plzf is expressed throughout the developing vertebrate nervous system

Plzf has a dynamic expression pattern throughout the CNS during vertebrate embryonic development (Cook et al., 1995, Avantaggiato et al., 1995). In the mouse and chicken nervous system, expression at the neural plate stage is initially widespread and becomes progressively more restricted as development continues (Cook et al., 1995). The authors find that expression becomes segmental within the hindbrain and later remains at high levels within the rhombomere boundaries whilst being downregulated elsewhere in the hindbrain. The expression of *Plzf* at rhombomere boundaries has also been observed in the rat CNS (Takahashi and Osumi, 2011).

The early expression pattern of *Plzf* in the murine hindbrain appeared to overlap with that of *Hoxb2*, suggesting that Plzf may play a role in regulating its expression. A Plzf binding site was located within the known rhombomere 3 / rhombomere 5 enhancer for murine *Hoxb2* (Sham et al., 1993, Ivins et al., 2003). Plzf was found to cooperatively bind to both this site and an A/T rich site known as 'Box1' and repress *Hoxb2* transcription (Ivins et al., 2003).

During neuronal differentiation of human pluripotent stem cells, neural progenitors spontaneously organise into structures termed 'neural rosettes' (Perrier et al., 2004). These neural cells have been isolated and maintained in culture where they were found to display high levels of Plzf expression (Elkabetz et al., 2008). The cells exhibit a broad differentiation potential and are believed to represent cells of the neural plate stage in embryos (Conti and Cattaneo, 2010). Recent experiments found Plzf to be expressed in neural epithelium cells of the developing human hindbrain, which could also form rosette-like structures when placed in culture (Tailor et al., 2013).

Studies of cultured human neuronal cells have shown that Plzf exerts a neuroprotective effect, acting along with the renin/prorenin receptor to mediate cell survival (Seidel et al., 2010). This work also examined expression of Plzf in the adult rat brain, finding expression to be ubiquitous throughout different brain regions and downregulated upon induction of an *in vivo* stroke model.

Whilst my PhD project was underway, work was published regarding a role of Plzf within neural progenitors of the chicken and mouse spinal cord (Gaber et al., 2013). They show Plzf is widely expressed throughout progenitors at the neural plate stage and later becomes restricted to the centre of the spinal cord where expression overlaps with FGF receptor 3 (FGFR3). Loss of Plzf was found to result in a slight increase in neurogenic markers and accompanied by reduced expression of genes associated with progenitor maintenance, suggesting that Plzf acts to maintain neural progenitors.

They find that ectopic Plzf results in increased FGFR3 expression, although the direct transcriptional target is not identified. This provides the cells of the central domain with a heightened responsiveness to FGF stimulation that, in turn, stimulates the STAT3 pathway. The authors therefore propose that the mechanism by which Plzf is acting is to modulate the ability of neural progenitors to respond to FGF ligands through downstream signalling pathways. The results and conclusions of their work will be discussed later in the context of my findings.

Role of Plzf within primary neurogenesis in zebrafish

Plzf became interesting to the Wilkinson lab during research into the function of the adaptor protein Btbd6a in zebrafish (Sobieszczuk et al., 2010). *btbd6a* was found to be expressed downstream of *neurog1* and knocking down expression resulted in a significant decrease in primary neurogenesis. It was hypothesised that Btbd6a was acting to target specific proteins for ubiquitination and degradation with the ubiquitin ligase protein Cul3 and therefore a proteome-wide search for interacting partners was carried out. This revealed Plzf as a potential binding partner. In the zebrafish this gene exists as two paralogues, *plzfa* and *plzfb*, arisen due to a genome duplication event.

Sobieszczuk et al. show that *plzfa* is highly expressed within the neural epithelium and that overexpression led to a significant decrease in the number of primary neurons produced, providing evidence that Plzfa is the aforementioned inhibitor. Plzfa knockdown did not lead to the increase in primary neurogenesis that would be expected from removing an inhibitor, but this was potentially explained by Notch-mediated lateral inhibition masking this effect. To test this, Notch signalling

was partially blocked through drug treatment and now when Plzfa was knocked down a significant increase in neuronal differentiation was observed. Knocking down Plzfa alongside Btbd6a revealed a complete rescue of the neurogenic phenotype, providing confirmation that removing Plzfa function is essential for neurogenesis to proceed.

Plzfa was characterised as acting to inhibit primary neurogenesis in a manner that is independent of Notch-mediated lateral inhibition. Further evidence showed that Btbd6a was capable of promoting nuclear export and degradation of Plzfa, thereby acting to keep levels of Plzfa low in cells selected to differentiate. Taken together, these results point towards the existence of a novel positive feedback loop whereby expression of proneural gene *neurog1* upregulates expression of *btbd6a* which targets Plzf for degradation (Figure 1-8).

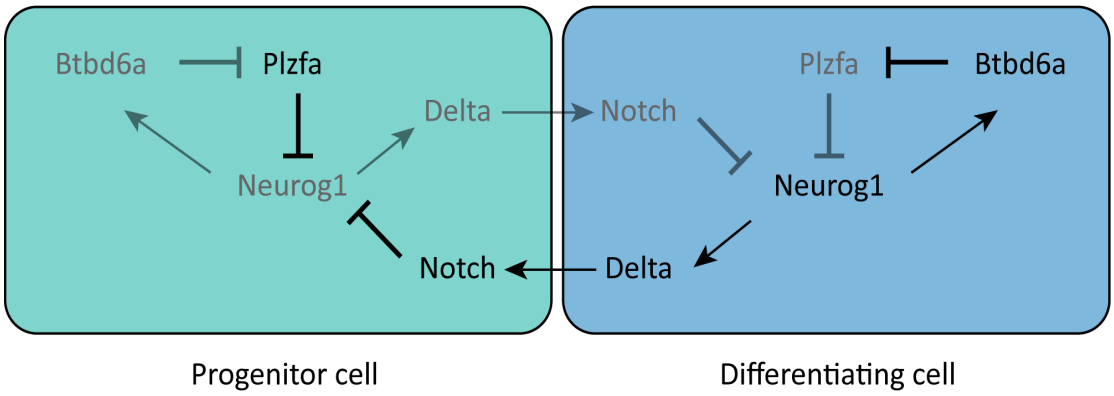
Figure 1-8: Role of Plzf in inhibiting primary neurogenesis

Plzfa acts to inhibit primary neurogenesis within the zebrafish. Cells selected to differentiate increase expression of *neurog1*, which in turn upregulates *btbd6a* expression. Btbd6a removes Plzfa from the nucleus, attenuating its inhibitory activity on Neurog1 activity. This feedback loop is independent and redundant to Notch-mediated lateral inhibition.

Figure adapted from Sobieszczuk et al. (2010).

Figure 1-8

Role of Plzfa in zebrafish primary neurogenesis



1.4 Aims of this study

The purpose of my PhD work is to further investigate the role that Plzf has within neural progenitors, using zebrafish as the model organism. Knowledge of the protein's function in other tissues leads to the hypothesis that Plzf is involved in the maintenance of neural progenitors. Plzf may function in spermatogonial and myeloid cells by regulating proliferation and differentiation of progenitors and therefore I aimed to determine whether Plzf is having a similar role in the nervous system.

Plzf is known to directly regulate transcription of target genes in various tissues, but no targets have been identified within the CNS. My first aim therefore was to determine transcriptional targets of Plzf within the developing zebrafish and compare this with my functional analysis in order to gain an appreciation of how Plzf is functioning within neural progenitors.

The previous work from the Wilkinson lab focused on the role of Plzfa during primary neurogenesis (Sobieszczuk et al., 2010). I therefore aimed to expand this analysis by investigating potential functional redundancy between Plzfa and its highly similar paralogue, Plzfb. Secondly, I sought to investigate whether Plzf functions during later stages of neurogenesis in the developing embryo.

Whilst my PhD project was underway, techniques to specifically modify the zebrafish genome were established. I therefore aimed to adopt this technology in order to assist with functional analysis of Plzf function.

2 Materials and Methods

2.1 Fish maintenance

Wild type and transgenic zebrafish embryos were obtained by natural spawning and maintained according to standard laboratory conditions (Westerfield, 1993). The *tp53* mutant line (Berghmans et al., 2005) was acquired from the Zebrafish International Resource Centre (University of Oregon, Eugene, USA). The *cldnb:lyngfp* transgenic line was used as described previously (Breau et al., 2012). Embryos were staged according to hours post fertilisation (hpf) and morphological criteria such as somite stage (ss) (Kimmel et al., 1995).

2.2 Zebrafish *in situ* hybridisation and immunohistochemistry

Embryos were grown to the desired stage and fixed in 4% paraformaldehyde / phosphate buffered saline (PBS) at 4°C overnight or at room temperature for 3 hours. Embryos that were fixed at time points over 30 hpf had been treated at 24 hpf with 0.2 mM phenylthiocarbamide (Sigma) to supress melanin synthesis. Fixed embryos were either stored in 100% methanol, or processed immediately for *in situ* hybridisation (ISH) or immunohistochemistry (IHC).

ISH probe synthesis

Previously described antisense RNA probes used were:
neurogenin1 (Gonzalez-Quevedo et al., 2010)
sox3 (Gonzalez-Quevedo et al., 2010)
neurod4 (Gonzalez-Quevedo et al., 2010)
btbd6a (Sobieszczuk et al., 2010)

For *plzfa,* PCR product against full length cDNA clone (Sobieszczuk et al., 2010) using AGAAGATGACGAGGAGCGG and TTGCCACATAGCTCGCATC primers. For *plzfb,* PCR product against full length cDNA clone (Sobieszczuk et al., 2010) using AAGACCGCAGGATCAAGTACC and TCACAGCCAAAGGTCTTCACTC.

The probes were synthesized using T3, T7 or SP6 RNA polymerase (Promega) and labeled either with Digoxigenin-UTP or fluorescein-UTP (Roche).

ISH reagents

Reagent	Components
Hybridisation Buffer	50% Formamide, 5x SSC, 50 µg/ml Heparin, 500 µg/ml tRNA, 5% Dextran Sulphate, 0.09 M Citric Acid (pH 6.0), 0.1% Tween
Washing Solution	50 % Formamide, 1x SSC, 0.1% Tween
MABT	20 mM maleic acid, 150 mM NaCl, 0.1% Tween, pH 7.5
Blocking Solution	20% Sheep Serum, 2% Roche Blocking Reagent in MABT
Staining Buffer	100 mM Tris-HCl pH 9.5*, 50 mM $MgCl_2$, 100 mM NaCl, 0.1% Tween
NBT/BCIP Substrate Solution	4.5 µl/ml NBT (Nitro blue tetrazolium, 75 mg/ml in 70% dimethyl formamide; Roche), 3.5 µl/ml BCIP (5-Bromo-4-chloro-3-indolyl phosphate, 50mg/ml in 70% dimethyl formamide; Roche) in Staining Buffer
Fast Blue Substrate Solution	2.5 µl/ml Fast Blue BB (100 mg/ml in dimethyl formamide; Sigma), 2.5 µl/ml NAMP (3-hydroxy-2-naphthoic acid 2,4-dimethylanilide phosphate, 100 mg/ml in dimethyl sulfoxide; Sigma) in Staining Buffer
Fast Red Substrate Solution	1 Fast Red/NAMP (Sigma) tablet in 2 ml 100 mM Tris-HCl pH 8.2, 0.1% Tween. Filtered before use.

Table 2-1: List of reagents used for ISH.
Asterisk (*) indicates 100 mM Tris-HCl pH 8.2 was used for samples developed with Fast Blue Substrate Solution

ISH protocol

Embryos were rehydrated through a graded series of methanol / PBS + 0.1% Tween-20 (Sigma) treatments (75%, 50%, 25%). Embryos were subsequently washed 5 x 5 minutes with PBST (PBS, 0.1% tween; Sigma). Embryos older than 24 hpf were treated with Proteinase K (10 µg/ml; Roche) in order to permeablise the embryo. The duration of the Proteinase K treatment depended upon the stage of the embryo and is described in Table 2-2. Following Proteinase K treatment, the embryos were washed 5 x 5 minutes with PBST and fixed for 20 minutes in 4% paraformaldehyde / PBS. Embryos were again washed 5 x 5 minutes with PBST before being left in Hybridisation Buffer at 65°C for 2 hours. Labelled riboprobes were added to the embryos and incubated at 65°C overnight.

The following day embryos were washed for 2 x 30 minutes in Washing Solution, 1 x 10 minutes in 1:1 Washing Solution:MABT, 3 x 5 minutes in MABT at 65°C, followed by 2 x 15 minutes with MABT at room temperature. Embryos were incubated in Blocking Solution for 2 hours. Anti-digoxigenin or anti-fluorescein

conjugated to alkaline phosphatase (1:1200; Roche) in 10% sheep serum was added overnight at 4°C. If required, primary antibodies for IHC were added at this point.

On the third day, embryos were washed 6 x 1 hour at room temperature with MABT and left overnight at 4 C. In order to develop the ISH signal, embryos were first washed 2 x 15 minutes at room temperature in Staining Buffer. The signal was developed in the dark by adding NBT/BCIP Substrate Solution. Development of the staining was monitored periodically and, once at the desired level, was stopped by washing 3 x 5 minutes with PBST, fixed at room temperature for 20 minutes in 4% PFA / PBS and washed a further 3 x 5 minutes before being processed for mounting as described below.

Developmental Stage (hpf)	Proteinase K treatment (minutes)
24	8
36	20
42	25
48	30

Table 2-2: Proteinase K treatment

Two colour fluorescent ISH

Two colour fluorescent ISH was carried out using the Fast Blue / Fast Red detection system previously described (Lauter et al., 2011). The fluorescein and digoxigenin labelled probes were added together to the embryos and incubated overnight. The protocol then proceeded as described above until the signal was developed. One of the probes was developed in the dark using the Fast Blue Substrate Solution until the blue staining was visible. At this point the embryos were washed 4 x 5 minutes with PBST followed by 3 x 20 minutes with 0.1 M glycine pH 3.0 to inactivate the antibody. The embryos were washed again 3 x 5 minutes with PBST, fixed for 20 minutes in 4% PFA / PBS at room temperature and washed 3 x 5 minutes with PBST. The anti-digoxigenin or anti-fluorescein alkaline phosphatase antibodies for the other probe were added in 10% sheep serum and incubated at 4°C overnight.

The next day the embryos were washed 10 x 15 minute with MABT at room temperature. In order to develop the Fast Red signal, embryos were incubated in 0.1 M Tris-HCL pH 8.2, 0.1% Tween for 3 x 10 minutes. The Fast Red Substrate

Solution was added and the signal was developed at room temperature in the dark. Development of the visible red signal was stopped and embryos were mounted as described below.

Detection of the fluorescent Fast Blue signal was carried out by excitation with the 633 nm laser and detecting wavelengths greater than 650 nm. The Fast Red signal was detected by excitation with the 561 nm laser and detecting wavelengths greater than 570 nm. Care was taken to not develop the Fast Blue signal for too long because the method involves the production of a precipitate that can obscure the subsequent of development of the Fast Red signal, making the detection of overlapping expression patterns difficult.

Immunohistochemistry antibodies

The primary antibodies used for immunohistochemistry (IHC) are listed in Table 2-3. An antibody was raised against the zebrafish Plzfa protein (zfPlzf Ab) from a Plzfa BTB domain peptide purified by Mohammed Ismail (unpublished). This rabbit polyclonal antibody was produced by Harlan Bioproducts.

Detection of primary antibodies for IHC was carried out using Alexa 488, 594 or 647 goat anti-rat, anti-rabbit, anti-mouse, anti-mouse IgG1 or anti-mouse IgG2b conjugates (1:500, Molecular Probes).

Immunohistochemistry protocol

Embryos were rehydrated as described in the ISH protocol. The samples were washed 5 x 5 minutes with PBST and blocked in 10% goat serum for 1 hour at room temperature. Primary antibodies (Table 2-3) were added in 10% goat serum and incubated at 4°C overnight. The following day the embryos were washed 6 x 30 minutes with PBST at room temperature. Secondary antibodies were added in 5% goat serum and incubated at 4°C overnight. On the final day, embryos were washed 4 x 30 minutes with PBST at room temperature, incubated with 4',6-diamidino-2-phenylindole (DAPI) for 30 minutes at room temperature, and finally washed 4 x 30 minutes with PBST before being processed for mounting.

Antigen (clone)	Species	Dilution	Source (Cat. No.)
Myc (9E10)	Mouse IgG_1	1:500	Santa Cruz (sc-40)
HA (3F10)	Rat IgG_1	1:200	Roche (1867423)
FLAG (M2) *	Mouse IgG_1	1:200	Sigma (F1804)
HuC/D (16A11)	Mouse IgG_{2b}	1:200	Molecular Probes (A-21272)
Zebrafish Plzf (zfPlzf)	Rabbit IgG (polyclonal)	1:2000	Wilkinson Lab
Mouse Plzf *	Mouse IgG_{2a}	1:200	Active Motif (39988)
Human Plzf *	Mouse IgG_{2a}	1:200	Calbiochem (OP128)
GFAP (6F2)	Rabbit IgG (polyclonal)	1:200	Dako (M076101-2)
zrf-1	Mouse IgG_1	1:200	Zebrafish International Resource Centre, University of Oregon
Sox2	Rabbit IgG (polyclonal)	1:200	Abcam (ab97959)
GFP	Rabbit IgG (polyclonal)	1:200	Torrey Pines Biolabs (TP401)
Prox1	Mouse IgG_1	1:100	Abcam (ab33219)
EphA4	Rabbit	1:450	Wilkinson Lab
Islet1/2 (39.4DS)	Mouse IgG_{2b}	1:100	Developmental Studies Hybridoma Bank, University of Iowa

Table 2-3: List of primary antibodies used for IHC.
Antibodies indicated with an asterisk (*) weren't detected in zebrafish IHC.

Mounting, imaging and processing

For transverse sections, 5-10 embryos in PBST were placed in a small plastic mould. The majority of the PBST was removed and melted 4% agarose / water was added to the embryos. Using forceps, the embryos were arranged in the required orientation before the agarose sets. The solid agarose blocks were kept under damp conditions overnight at 4°C. The embryos were sectioned using a Vibratome (Lecia VT1000 S), generating transverse sections of a thickness between 80 – 120 µm.

For all other imaging purposes, embryos were stored in 70% glycerol / PBS and kept at 4°C until required. Prior to mounting the yolk was manually removed from the embryo. For flat mounting, embryos were placed with their ventral surface contacting a glass slide. For side mounting, embryos were placed with a lateral surface contacting the slide. A coverslip was placed onto the samples such that it is in contact with the surface of the embryo and 70 % glycerol added.

Samples were visualised using a Zeiss Axio Imager.Z2 microscope fitted with a Zeiss Axiocam digital camera. When specified, fluorescent images are captured using a Leica TCS SP2 confocal microscope. Images were processed using ImageJ (NIH) and Adobe Photoshop.

2.3 Morpholino oligonucleotides and RNA microinjections

Morpholinos

Morpholino oligonucleotides (MO) were purchased from Gene Tools (Oregon, USA) and dissolved, aliquoted and stored at room temperature at a concentration of 1 mM as previously described (Gerety and Wilkinson, 2011). Blastomeres (1- to 4-cell) were microinjected with 5 – 10 ng MO as described in text. Unless stated otherwise, MOs were injected into p53 null mutant embryos to avoid MO mediated off-target toxicity (Robu et al., 2007, Gerety and Wilkinson, 2011). The following translation blocking MO sequences were used:

plzfa MO, TCTCTGAAATCCACACGGCCAACTC

plzfb MO, ACATCAAGATTTACCGAACCATCTC

neurog1 MO, ATACGATCTCCATTGTTGATAACCT (Sobieszczuk et al., 2010)

p53 MO, GCGCCATTGCTTTGCAAGAATTG (Robu et al., 2007)

Control MO, CCTCTTACCTCAGTTACAATTTATA

Capped RNA

Capped RNA was synthesised using the SP6 mMessage mMachine kit (Ambion) from linearised DNA. The RNA was purified by phenol:chloroform extraction and isopropanol precipitation as described in the manufacturer's protocol. Constructs used to generate capped RNA, other than those generating TALEN RNA, are described below in Table 2-4. Between 30 – 300 pg was injected into the one cell stage of the embryo, as specified in the text.

Construct	Source
pCS2-H2B-Citrine	Megason Lab
pCS2-H2B-RFP	Megason Lab
pCS2-HA-Plzfa	Wilkinson Lab (Sobieszczuk et al., 2010)
pCS2-Myc-Plzfa	Wilkinson Lab (Sobieszczuk et al., 2010)
pCS2-Plzfa-FLAG	Wilkinson Lab
pCS2-Plzfa-Myc	Described below
pCS2-Myc-Btbd6a	Wilkinson Lab (Sobieszczuk et al., 2010)
pCS2-Neurog1	Wilkinson Lab (Sobieszczuk et al., 2010)

Table 2-4: List of constructs used for generating capped RNA

2.4 Cell culture, transfection and immunohistochemistry

HEK293 cells were grown as described (Poliakov et al., 2008). Cells were transfected with 1 ug of the appropriate plasmid using FuGENE HD Transfection Reagent (Promega) according to manufacturers instructions. For IHC, cells were fixed after 48 hour using 4% formaldehyde for 15 min at room temperature, washed with PBST, blocked for 1 hour in 4% donkey serum and 2% bovine serum albumin and stained using primary and subsequently secondary antibodies diluted in blocking buffer (Odyssey).

2.5 Design and construction of Transcription Activator-Like Effector Nucleases (TALENs)

TALENs were designed and built using principles outlined previously (Cermak et al., 2011, Huang et al., 2011, Sander et al., 2011, Dahlem et al., 2012). Briefly, each TALEN array was designed such that 1) it recognised 16 – 20 base pairs of sequence, 2) a spacer of length 14 – 17 bp falls between the two arrays in the targeting pair, 3) a thymine base is immediately upstream of the DNA targeted by each array and 4) no homology with other genes was found for the region targeted. A list of potential TALENs was generated using the TALEN targeter software available online (https://boglab.plp.iastate.edu/ (Cermak et al., 2011)) and subsequent selection was performed manually.

TALEN construction was performed using the Golden Gate cloning technique designed for rapid generation of large constructs (Engler et al., 2009, Cermak et al., 2011). The set of plasmids to do this was obtained from Addgene (Cat #1000000016) and the designed TALENs were built using the 5-day protocol

described (Cermak et al., 2011). Golden Gate compatible destination vectors pCS2TAL3-DD and pCS2TAL3-RR (Dahlem et al., 2012) were obtained from Addgene (Cat. No. 37275 and 37276). Once constructed, the DNA was linearised using the NotI restriction enzyme and capped RNA synthesised using the SP6 mMessage Machine kit (Ambion). Equal amounts of left and right TALEN mRNA were injected together into the one cell stage zebrafish embryo.

Plasmids encoding heterodimeric FokI containing the KK and EL mutations described previously (Miller et al., 2007) were obtained from Addgene (Cat. No. 21872 and 21873). The FokI was excised from these fragments individually and cloned into pCS2TAL3-DD and pCS2TAL3-RR respectively, forming Golden Gate compatible destination vectors pCS2TAL3-KK and pCS2TAL3-EL. Table 2-5 contains a list of the various TALEN arrays constructed.

Construct Name	Gene targeted	FokI backbone
***golden* TALEN Left**	*slc24a5 (golden)*	pCS2TAL3-DD
***golden* TALEN Right**		pCS2TAL3-RR
***plzfa* TALEN Left**	*plzfa*	pCS2TAL3-DD
***plzfa* TALEN Right**		pCS2TAL3-RR
***plzfa* TALEN Left (KK)**		pCS2TAL3-KK
***plzfa* TALEN Right (EL)**		pCS2TAL3-EL
***plzfb* TALEN Left**	*plzfb*	pCS2TAL3-DD
***plzfb* TALEN Right**		pCS2TAL3-RR
***redd1* TALEN Left**	*redd1*	pCS2TAL3-DD
***redd1* TALEN Right**		pCS2TAL3-RR
***btbd6a* TALEN Left**	*btbd6a*	pCS2TAL3-DD
***btbd6a* TALEN Right**		pCS2TAL3-RR

Table 2-5: List of TALEN constructs

2.6 Generating nickases

Site directed mutagenesis was carried out using the QuikChange II kit (Stratagene) following manufactures' instructions. The previously described primers (Kim et al., 2012) used to generate the D450A mutation in the *FokI* domain were:

D450A Fwd: GATCAAGGAAACCGGCCGGAGCAATTTATAC

D450A Rv: GTATAAATTGCTCCGGCCGGTTTCCTTGATC.

2.7 Genotyping

Genomic DNA (gDNA) was obtained from individual three days post fertilisation embryos or from fin clips from adult fish by lysis in 50 ul DNA extraction buffer (10 mM Tris-HCl (pH 8.0), 10 mM EDTA, 0.2% Triton, 0.2 mg/ml Proteinase K) at 55°C for 9 hours. The reaction was terminated by heating the samples at 95°C for 20 minutes.

The presence of TALEN-induced mutations was determined by High Resolution Melt (HRM) curve analysis, as previously described (Dahlem et al., 2012). Primers were designed using Primer3 (http://frodo.wi.mit.edu/primer3/) to amplify ~ 100 bp gDNA around the TALEN pair target site. Triplicate 20 µl reactions containing 1 µl of gDNA were amplified and denatured in the presence of the MeltDoctor HRM Dye (Applied Biosystems) using an Applied Biosystems 7900HT Fast Real-Time PCR System according to manufacturers instructions. HRM data was analysed using the Applied Biosystems HRM Software v2.0 and used to generate the melt profiles shown within the text. Minor image improvements and addition of annotations were performed using Adobe Illustrator.

When specified, around 500 bp of gDNA was amplified by the polymerase chain reaction (PCR) and subject to downstream analysis. For sequencing individual alleles, the PCR product was cloned into pGEM-T Easy (Promega), transformed into competent bacteria, and individual colonies were picked to carry out colony PCR using the SP6 and T7 primers. The resulting PCR product was sequenced (GATC Biotech). Restriction Fragment Length Polymorphism (RFLP) was carried out in order to analyse the loss of a restriction enzyme recognition site present within the TALEN spacer region. The PCR products were digested with the restriction enzymes described within the text and resolved on a 2% agarose gel.

Name	Sequence	Purpose
***golden* HRM Fwd**	CAGGAGAGGAAAGATGGAGGAA *	HRM
***golden* HRM Rv**	GCTGATGACCTCCAGAGATGG *	HRM
***plzfa* Target Site HRM Fwd**	GCCGTGTGGATTTCAGAGAC	HRM
***plzfa* Target Site HRM Rv**	GCGCATCTGATTAGCCTTGT	HRM
***plzfa* Off-Target Site HRM Fwd**	GGAATGATCCAGCTCCAGAA	HRM
***plzfa* Off-Target Site HRM Rv**	CCACCATGATGACCACGTC	HRM
***plzfb* Target Site HRM Fwd**	GTTCTGTGCGCATGAAACTC	HRM
***plzfb* Target Site HRM Rv**	CAGCCACCCTACAACTCTCC	HRM
***plzfb* Off-Target Site HRM Fwd**	GATGGTTCGGTAAATCTTGATG	HRM
***plzfb* Off-Target Site HRM Rv**	AGACAGGCGCATCTGGTT	HRM
***btbd6a* HRM Fwd**	TGGCTACACTCTATGCTGCAA	HRM
***btbd6a* HRM Rv**	ATTTCGCGCCTCCAGACT	HRM
***redd1* HRM Fwd**	CGTCCACACCGACATCTG	HRM
***redd1* HRM Rv**	TGTGGTTGTCGGAGTCAGAG	HRM
***plzfa* PCR Fwd**	ACAAGAAAACGAACAACTGCAA	PCR / RFLP
***plzfa* PCR Rv**	CTTGGAGCGTGGCAGTGTAG	PCR / RFLP
***plzfb* PCR Fwd**	CAGTTGCAGGAGCACTCAAG	PCR / RFLP
***plzfb* PCR Rv**	AACCGCCATCTTGTATGGAA	PCR / RFLP
F1	GCCGTGTGGATTTCAGAGAC	PCR
F2	GAAATCGGCAAAATCCCTTA	PCR
F3	CAGGAGAGGAAAGATGGAGGAA *	PCR
F4	TGAGTTTGGACAAACCACAAC	PCR
R1	TGAAGTCGATGCCCTTCAG	PCR
R2	ACTCCTGGCTGTCCACCAT	PCR
R3	CTGTTCCGGTGAAAGAGGAT	PCR
SP6	ATTTAGGTGACACTATAGAA	Colony PCR
T7	TAATACGACTCACTATAGG	Colony PCR

Table 2-6: List of primers used for genotyping
Golden HRM primers (*) were previously described (Dahlem et al., 2012).

2.8 Targeted knock-ins

Construction of donor plasmids

The non-expression vector pBluescript II KS was used as a backbone for donor plasmid construction. Following guidelines (Maresca et al., 2013) the left and right TALEN binding sites at the *plzfa* locus were inverted whilst the spacer region remained in the same direction as in the zebrafish genome. The coding region for eGFP from eGFP-N1 (Clontech) is preceded by the sequence encoding the 19 amino acid PTV1-2A (P2A) motif (Poulain and Ober, 2011) and followed by the SV40 polyadenylation signal from the $pCS2^{+}$ vector. Flanking the 2A-eGFP-polyA cassette are two 39 bp attP sites (Hu et al., 2011). The entire 1271 bp region was synthesised

by GENEWIZ and subcloned into pBluescript II KS, generating *plzfa* eGFP Donor Plasmid. A plasmid map for the *plzfa* eGFP Donor Plasmid can be found in Appendix Figure 8-6. For other donor constructs, eGFP was removed and H2B-citrine from pCS2-H2B-citrine was cloned in-frame to generate P2A-H2B-citrine-polyA (*plzfa* Citrine Donor Plasmid). The GAL4; UAS-tdTomato cassette was removed from a Tol2-CMLC2-GAL4; UAS-tdTomato plasmid (Gerety S, unpublished) and used to replace the eGFP cassette, generating P2A-GAL4-polyA; UAS-tdTomato-polyA (*plzfa* GAL4 Donor Plasmid). The *alpha crystallin* promoter (Kurita et al., 2003) driving red fluorescent protein (RFP) or citrine was taken from previously described constructs (Gerety and Wilkinson, 2011) and inserted *3'* to the expression cassettes. Sequences for the described cassettes can be found in Appendix Figure 8-8. For *plzfa* Donor Plasmid v2, a second inverted TALEN binding site was synthesised by GENEWIZ and subcloned into the *plzfa* eGFP Donor Plasmid, *3'* to the eGFP cassette.

In order to construct a donor plasmid targeting the *golden* gene, the inverted *plzfa* TALEN binding sites were excised by digesting the plasmid with the restriction enzymes *SacI* and *ClaI*. A pair of short oligonucleotides containing the *golden* TALEN binding sites were designed such that when annealed they would contain overhangs complementary to *SacI* and *ClaI* as shown below. The oligonucleotides were annealed following a method described previously (Hwang et al., 2013b). Briefly, Annealing Buffer (40 mM Tris pH 8.0, 2 mM $MgCl_2$, 50 mM NaCl, 1 mM EDTA pH 8.0) and was added to 100 µM of each oligonucleotide. Using a thermocycler the sample was heated to 95 ^{o}C and decreased by 1°C every 30 seconds until the sample reached 4°C. 3 µl of the annealed oligonucleotide product was added to 1 µl digested plasmid (5 ng/µl) in a 10 µl ligation reaction. Competent cells were transformed and bacteria grown using standard cloning procedures.

golden Forward Oligonucleotide: *5'* CAGATAGATCTTGGCAGAAAATATTCATCTCCATCG TGTGTGACGGACAGCAGCATGTAGTAT *3'*
golden Reverse Oligonucleotide: *5'* TCGAGTCTATCTAGAACCGTCTTTTATAAGTAGAGG TAGCACACACTGCCTGTC GTCGTACATCATAGC *3'*

Injection and Screening for Insertion

9 pg to 99 pg of donor plasmid with or without mRNA encoding the appropriate TALEN pairs was injected into the cell of 1-cell stage zebrafish embryos as specified in the text. Numbers of dead, deformed (monsters) and morphologically wild-type embryos were scored at 24 hpf. Between 24 and 72 hpf, embryos were screened and scored for expression of fluorescent proteins by observing embryos using a Leica M205FA stereoscope. Embryos older than 48 hpf were anesthetized with 0.02% 3-aminobenzoic acid ethyl ester (MS222) in order to restrict their movement before screening. When noted in the text, GFP expression was detected by performing IHC with anti-GFP, otherwise fluorescence was detected without adding antibodies. Analysis by PCR was carried out using primers described in Table 2-6 and resolved on a 1.5% agarose gel.

2.9 Chromatin Immunoprecipitation

Chromatin immunoprecipitation antibodies and reagents

The antibodies used for chromatin immunoprecipitation (ChIP) are listed in Table 2-7. Details of their use can be found within the main text. The recipes for the various buffers used throughout the ChIP process are shown in Table 2-8. 1 tablet of protease inhibitor cocktail tablets (Roche) and 20 µl protease inhibitor cocktail (Sigma) were added to 10 ml of the specified buffers shortly before their use.

Antigen	Species	Source (Cat. No.)
Myc	Goat	Abcam (ab9132)
Zebrafish Plzf	Rabbit	Wilkinson Lab
Mouse Plzf	Mouse IgG_{2a}	Active Motif (39988)
Human Plzf	Mouse IgG_{2a}	Calbiochem (OP128)
H2A.Z	Rabbit	Abcam (ab4174)

Table 2-7: List of antibodies used for ChIP.

Reagent	Components
Cell Lysis Buffer *	10 mM Tris-HCl (pH 8.1), 10 mM NaCl, 0.5% NP-40
Nuclear Lysis Buffer *	50 mM Tris-HCl (pH 8.1), 10 mM EDTA, 1% Sodium Dodecyl Sulphate (SDS)
ChIP Dilution Buffer * †	16.7 mM Tris-HCl (pH 8.1), 167 mM NaCl, 1.2 mM EDTA, 0.01% SDS, 1.1% Triton
Low Salt Wash Buffer †	20 mM Tris-HCl (pH 8.1), 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton
High Salt Wash Buffer †	20 mM Tris-HCl (pH 8.1), 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton
LiCl Wash Buffer †	10 mM Tris-HCl (pH 8.1), 0.25 M LiCl, 1 mM EDTA, 1% IGEPAL CA630, 1% deoxycholic acid (sodium salt)
TE †	10 mM Tris-HCl (pH 8.0), 1 mM EDTA
Elution Buffer	50 mM $NaHCO_3$, 1% SDS

Table 2-8: List of reagents used for ChIP.
Buffers marked by an asterisk (*) had protease inhibitors added just before use. Buffers marked with a dagger (†) were purchased from Millipore. All buffers were kept at 4°C and filtered before use.

ChIP Protocol

The zebrafish ChIP protocol was adapted from Junji Lin's protocol from Richard Dorsky's Lab (University of Utah) found on the ZFIN protocols web page (https://wiki.zfin.org/display/prot/Chromatin+Immunoprecipitation+%28ChIP%29+Protocol+using+Dynabeads). Significant assistance in developing the protocol was provided from Emmanouil Metzakopian and Ben Martynoga (NIMR).

Deyolking and cross-linking

Zebrafish embryos were grown to either 12 hpf or 24 hpf as specified in the text. 100 embryos at a time were manually dechorionated in water using forceps and transferred to a Non-stick RNase-free 1.5 ml tube (Ambion). For all following stages the tubes were kept on ice as much as possible. The water was replaced with 150 µl ice cold Hanks Buffered Saline Solution (Magnesium and Calcium free; Invitrogen). Deyolking was performed by mechanical stress caused by pipetting embryos up and down 5 – 8 times using a 200 ul tip. Deyolked embryos were pooled together and embryos were pelleted by spinning at 300 g for 30 seconds at 4°C. The

solution was removed and the embryos were washed with 1 ml Hanks Buffered Saline Solution, spun again, and replaced with 1 ml ice cold PBS.

To cross-link the protein and DNA, 27 µl formaldehyde (37% stock; Sigma) was added drop wise to the PBS, making a final concentration of 1%. Tubes were rotated for 15 minutes at room temperature. 50 µl 2.5 M glycine was added in order to quench the formaldehyde and the tubes were rotated for a further 5 minutes at room temperature. The embryos were pelleted by spinning at 300 g for 30 seconds at 4°C and washed twice using 1 ml ice cold PBS.

Embryo lysis and chromatin extraction

The pellet was resuspended in 1 ml Cell Lysis Buffer (for 300 embryos). The tubes were left on ice for 20 minutes, pipetting up and down every 5 minutes until no clear tissue was visible. Tubes were spun at 1200 g for 5 mintues at 4°C, washed with 1 ml Cell Lysis Buffer and spun again. The white nuclear pellet was resuspended with 250 µl Nuclear Lysis Buffer (per 100 embryos). Tube were left laying on ice for 20 minutes before being flash frozen in liquid nitrogen and stored at -80°C. Chromatin was stored like this for up to a month.

Sonication

Samples were thawed on ice and split into 1.5 ml tubes, so that each contains 250 µl chromatin. For fragments between 0 – 400 base pairs (bp), samples were sonicated using a Diagenode Bioruptor for a total of 40 mintues on high setting using 30 second intervals. To avoid potential denaturing of the protein, every 15 minutes the tubes were removed from the Bioruptor and left lying on ice for 10 minutes and the water within the Bioruptor was replaced with ice-cold water. For fragments between 0 – 1000 base pairs (bp), total sonication time was 10 minutes. Debris was removed by centrifugation at 13 000 g for 10 minutes at 4°C and the supernatant was collected from each tube and pooled together. A 50 µl sample was taken for reverse crosslinking (described below) to check sonication conditions. 1% of the total sonicated chromatin was put aside to be used as an input. This sample was stored at 4°C until the ChIP sample was ready to be reverse cross-linked.

Immunoprecipitation

Between 10 – 30 µg chromatin was used per ChIP as specified in the text. 50 µl Protein G Dynabeads (Invitrogen) per IP were equilibrated by washing 3 x with 1 ml ChIP Dilution Buffer and subsequently were resuspended in their original volume. In all cases the beads and supernatant were separated by a brief pulse centrifugation followed by placing the tubes on a magnetic separation stand (Promega). The equilibrated beads were added to the chromatin sample and the volume was made up to 1 ml with ChIP Dilution Buffer. The tubes were rotated for 2 hours at 4°C. The supernatant was separated and placed in a new 1.5 ml tube to which the specified amount of antibody was added and the samples were rotated overnight at 4°C. Control samples were rotated without the addition of any antibody. A separate aliquot of Protein G Dynabeads (50 µl / ChIP) were blocked by rotating the beads overnight at 4°C in the presence of 1 mg/ml bovine serum albumin in ChIP Dilution Buffer.

The next day the blocked beads were washed twice with 1 ml ChIP Dilution Buffer and resuspended in their original volume. 50 µl of these blocked beads were added to each chromatin sample and the tubes were rotated for 1 hour at 4°C to allow the antibody – chromatin mixture to bind to the beads. To remove non-specific binding, the samples were washed in different buffers by rotating at room temperature for 5 minutes, removing the supernatant after each wash. The order of the washes was 1 x Low Salt Wash Buffer, 1 x High Salt Wash Buffer, 1 x LiCl Wash Buffer, 2 x TE.

To extract the DNA from the beads, 250 µl freshly made Elution Buffer was added to the beads and samples incubated at 65°C for 10 minutes, vortexing the samples every 10 minutes. The supernatant was transferred to a new 1.5 ml tube and the elution process was repeated, generating a final 500 µl eluate.

For experiments where the protein was required from the ChIP experiment, NuPAGE LDS sample buffer (Invitrogen) with Reducing Agent was diluted 2x with water and 20 µl was added to the beads. Samples were incubated at 95°C for 10 minutes and the supernatant was transferred to a new 1.5 ml tube. Protein samples were briefly incubated on ice and then frozen at -80°C before being analysed on a Western Blot as described below.

Reverse cross-link and DNA Purification

Both the 1% input and the ChIP samples were made up to 500 µl using PBS. To all samples, 10 µl RNase A was added to a final concentration of 0.2 µg/µl. The samples were incubated at 37°C for 1 hour. 5 µl of Proteinase K (10 mg/ml) and 11 µl of 5 M NaCl were added to the samples and incubated at 42°C for 2 hours. Finally, all samples were incubated at 65°C overnight.

DNA was extracted with 500 µl phenol/chloroform by vortexing for 30 seconds followed by centrifuging at full speed for 3 minutes. The aqueous phase was retained and the sample extraction procedure was carried out using 500 µl chloroform. The aqueous phase was transferred to a new 1.5 ml tube and 1 µl glycogen, 50 µl 3 M NaAc pH 5.2 and 900 µl isopropanol was added to the samples. Samples were incubated at -20°C for 30 minutes followed by centrifugation at full speed for 30 minutes at 4°C. The resulting pellet was washed with 75% ethanol and left to air dry. Samples were resuspended in 35 µl water and kept at -20°C until required.

Quantitative PCR (qPCR)

ChIP-qPCRs were assembled using Platinum SYBR Green Super mix (Invitrogen). Reactions were performed in triplicates on an Applied Biosystems 7900HT Fast Real-Time PCR System according to manufacturers instructions. Enrichment was calculated by using the $2^{-\Delta\Delta Ct}$ method comparing Ct values for genomic regions obtained by ChIP using the specific antibody and ChIP using no antibody. Primers were designed using Primer3 software and tested using multiple dilutions of input genomic DNA and a dissociation curve was used to make sure a single product was generated.

Name	Sequence
***copb1* TSS Fwd**	GCTAACTCGAGCCCTGAATC *
***copb1* TSS Rv**	ACTGATTGCGTGAACAGTCG *
***copb1* Exon Fwd**	TCAGACAGATGTGGGCTGAG *
***copb1* Exon Rv**	TCCAGACATTCCCGCTAAAC *
***hsf2* TSS Fwd**	TGGTGCGTGTGTTTTGTCTAC *
***hsf2* TSS Rv**	CGAGCTGTGTTTCATTTTGC *
***hsf2* UTR Fwd**	CAGACCTGCTGGATGAGTCC *
***hsf2* UTR Rv**	GATAGAAGAGCGTGGCTTCG *
***klhl20* TSS Fwd**	AGTCCTCACGTCACACATGG *
***klhl20* TSS Rv**	ACACTGGCGAAAGGAGTGTC *
***klhl20* UTR Fwd**	AATGCATCAACGAGTCAGTCC *
***klhl20* UTR Rv**	ACCTCGGGTATCAAACTTGC *
***crabp1* Fwd**	ATTTCCCGTTTTTGCTGTCA
***crabp1* Rv**	AGCAGTAAGTCTAGCTCCAGCA
***ccna2* Fwd**	AAGTAGCCCGCGACTATTGA
***ccna2* Rv**	ACTCTCAACGGCTCGCTTT
***hoxd11* Fwd**	CAAAGAGAACCGCAAAAACG
***hoxd11* Rv**	TTGACTATCTTAAAATGTCGAATCAAA
***deltaA* Exon Fwd**	TGTCTGAATGGAGCCACTTG
***deltaA* Exon Rv**	GCATTCGTTGACCTCGATCT
***ephrinA1* Exon Fwd**	GTGCTCTACATGGTGGAACG
***ephrinA1* Exon Rv**	AAAACTTTTCGGGAGCATGA
***elavl3* Exon Fwd**	CGTGAGGTGATGATCCTTCC
***elavl3* Exon Rv**	CGATGCCAACCTGTATGTGA

Table 2-9: List of primers used for ChIP-qPCR.
Primers marked with an asterisk (*) were provided by George Gentsch (NIMR). TSS = Transcription Start Site, UTR = Untranslated Region

2.10 Western Blots

Embryos were deyolked as described above and lysed in 30 µl of Cell Lysis Buffer (with protease inhibitors) for 20 minutes on ice. Debris was removed by spinning at full speed for 10 minutes at 4°C and the supernatant collected. Protein concentration in the cell lysates was measured using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific), according to manufacturer's instructions. 1x NuPAGE LDS Sample Buffer (Invitrogen) was added along with reducing agent to the samples before denaturing at 95°C for 10 minutes. Polyacrylamide gel electrophoresis of proteins was carried out by loading > 30 µg of protein onto a NuPAGE 10% Bis-Tris gel (Invitrogen) with 5 µl Novex Sharp Pre-stained protein standard, and run in

Novex Mini-Cell tanks (Invitrogen) in NuPAGE MOPS SDS Running Buffer (Invitrogen) for 45 min at 200V. The entire 20 µl ChIP protein samples were loaded onto the gel when specified.

Transfer of proteins from gels on Immobilon-FL membranes (Milipore) was performed using XCell II Blot module (Invitrogen) under manufacturers instructions. Membranes were blocked in blocking solution (50% Odyssey Blocking Buffer (LI-COR Biosciences), 50% PBS) for 1 hour at room temperature and stained with the appropriate primary antibodies as described in Table 2-10, diluted in blocking solution for 1 hour at room temperature. This was followed by staining with secondary antibodies conjugated to infrared fluorescent dyes IR700 and IR800 (Rockland) diluted 1:5000 in blocking solution containing 0.1% Tween and 0.01% SDS. After staining, the membranes were scanned using 700 and 800 nm channels on an imager (Odyssey).

Antigen	Species	Dilution	Source (Cat. No.)
Myc (9E10)	Mouse IgG_1	1:1000	Santa Cruz (sc-40)
α-Tubulin	Mouse IgG_1	1:1000	Sigma (T9026)
γ-Tubulin	Rabbit IgG	1:1000	Sigma (T3559)
Zebrafish Plzf	Rabbit IgG (polyclonal)	1:400	Wilkinson Lab
HDAC1	Rabbit IgG (polyclonal)	1:3200	Abcam (ab7028)

Table 2-10: List of antibodies used for Western Blots.

2.11 High Throughput Sequencing

DNA from the ChIP (experimental) and input (control) DNA samples were quantified using the Qubit Fluorometer (Invitrogen). 10 ng of both samples were used to generate sequencing libraries following the manufacturer's protocol (Illumina). Quality control, library preparation and running the samples on a HiSeq 2000 (Illumina) was performed the NIMR High Throughput Sequencing department. 40 bp reads were obtained.

Data analysis was carried out using the publically available Galaxy server (https://usegalaxy.org/) (Goecks et al., 2010, Blankenberg et al., 2010, Giardine et al., 2005). The raw reads for the ChIP and input samples were mapped to the

zebrafish genome (zv9) with BowTie version 1.1.2 (Langmead et al., 2009). Peak calling was performed using Model-based Analysis of ChIP-Seq (MACS) version 1.0.1 (Zhang et al., 2008b). DNA binding events were visualised using the IGV browser (Robinson et al., 2011).

3 A genome-wide search for transcriptional targets of Plzf

3.1 Introduction

Understanding where in the genome transcription factors bind can be highly informative in dissecting their function. In the hematopoietic and spermatogonial stem cells, Plzf directly regulates genes associated with proliferation (Shaknovich et al., 1998, Doulatov et al., 2009), and differentiation (Doulatov et al., 2009, Filipponi et al., 2007, Hobbs et al., 2010), in order to maintain the progenitor state. Whilst a function for Plzf within the developing nervous system has been described (Sobieszczuk et al., 2010, Gaber et al., 2013), no direct transcriptional targets have been identified within this tissue. In order to further understand how Plzf regulates neurogenesis and possibly discover new functions, I aimed to perform a genome-wide characterisation of Plzf targets.

To identify genomic targets I planned on using the chromatin immunoprecipitation (ChIP) technique on zebrafish embryos followed by high-throughput sequencing (ChIP-Seq). This process allows sequencing of the genomic DNA fragments that are bound by a DNA-binding protein, such as a transcription factor. Once DNA fragments are sequenced, bioinformatics techniques are used to assign these regions to an organism's genome; theoretically identifying regions that the protein binds to *in vivo* in an unbiased manner (Liu et al., 2010).

In this chapter I will first of all evaluate the ChIP technique within zebrafish embryo using a known ChIP-grade antibody against the histone protein H2A.Z. I shall describe my efforts to determine whether I am able to use an antibody raised against the zebrafish Plzfa protein, zfPlzf Ab, for ChIP studies. Presenting evidence that zfPlzf Ab is unsuitable for ChIP, I elect to carry out ChIP-Seq against a tagged version of Plzfa after showing evidence that it can be immunoprecipitated from zebrafish chromatin. Finally, I show evidence that the ChIP-Seq experiment failed to identify regions bound by Plzf and troubleshoot the possible reasons why.

3.2 Evaluating the ChIP technique in zebrafish

There are several technical considerations to take into account when performing ChIP-Seq experiments. Firstly, I had to decide what to use as my source of chromatin. There are advantages to carrying out ChIP-Seq using cultured stem cells, such as working with a homogenous population of cells and the ease of obtaining large amounts of material to work with. However, binding events that occur *in vitro* may not directly correlate with events occurring *in vivo*. Therefore, I chose to perform ChIP on whole zebrafish embryos because my functional analysis is being performed in the same system and previous studies have demonstrated the ability to identify genomic targets within zebrafish (Wardle et al., 2006, Aday et al., 2011, Xu et al., 2012).

A major factor that contributes to the success of the ChIP-Seq experiment is the choice of antibody (Kidder et al., 2011). A frequently used method to determine whether an antibody is suitable for ChIP experiments is by detection of locus-specific enrichment at sites directly bound by the transcription factor of interest. This can be done by ChIP-PCR assays when the enrichment at several positive control regions is compared to that at negative control regions.

H2A.Z ChIP-qPCR

In order to determine the ideal experimental conditions for the ChIP technique in zebrafish, I decided to perform ChIP-PCR using a known ChIP-grade antibody that recognises the histone variant H2A.Z protein. H2A.Z has a variety of functions and is known to be associated with nucleosomes close to the transcriptional start site (TSS) of genes (Zlatanova and Thakar, 2008). Previous studies have show enrichment of H2A.Z at the TSS of several genes within the zebrafish genome by ChIP-PCR (George Gentsch, unpublished).

I initially chose to work with 24 hpf embryos because at this stage there are a high proportion of neural progenitors that are actively proliferating and undergoing neurogenesis. I therefore wanted to determine the conditions required to reversibly cross-link DNA and proteins and to fragment the DNA at this stage in development. The basic ChIP protocol is shown in Figure 3-1a and described in detail within Chapter 2.9. ChIP-Seq ideally requires around 5 x 10^7 cells starting material (Schmidt

et al., 2009, Bogdanović et al., 2013) and because a 24 hpf embryo is estimated to contain between 20,000 – 30,000 cells (Aday et al., 2011) I elected to extract chromatin from 500 embryos per experiment.

Sonication conditions were determined in order to fragment the cross-linked chromatin to range between 200 – 1000 bp, suitable for ChIP-PCR (Figure 3-1b). For initial PCR analysis, this range of DNA fragments should suffice, although further sonication such that I obtain fragments between 150 – 300 bp is required for ChIP-Seq. From the 500 24 hpf embryos, I obtained between 60 – 80 µg of DNA and performed individual ChIP experiments using 10 µg DNA. Samples were immunoprecipitated with either an antibody specific to H2A.Z or went through identical treatment without the addition of an antibody (negative control).

In order to confirm that the H2A.Z ChIP experiment has worked, the samples were analysed by ChIP-PCR (Figure 3-1c). Positive control regions were the TSS of selected genes (*copb1*, *hsf2* and *klhl20*) and negative control regions were locations within the exons and *3'* UTRs of the respective genes, several kilobases away from the TSS. Between 7 – 11 fold enrichment was observed at the positive control regions for the three genes tested when ChIP was carried out using the H2A.Z antibody compared to the no antibody control. Encouragingly, no enrichment was detected at the negative control regions. Overall, these results suggest that ChIP-PCR for H2A.Z using zebrafish embryos was successful and therefore I can use the same protocol to begin ChIP for Plzf.

Characterising the zfPlzf Ab

Ideally I sought to detect where the endogenous Plzf protein binds within the zebrafish genome. A caveat with this approach is that there is no known ChIP-grade antibody that works in zebrafish. Before I began my PhD project, the Wilkinson Lab raised a polyclonal antibody raised against the N-terminal BTB domain of the zebrafish Plzfa protein (zfPlzf Ab). It was important that this antibody was fully characterised since the aim was to use it for multiple purposes during my PhD work. As previously mentioned, there are two paralogues of Plzf within zebrafish, described as Plzfa and Plzfb. These proteins share over 90% similarity in their BTB domain and as the antibody is raised against numerous epitopes, it seemed likely

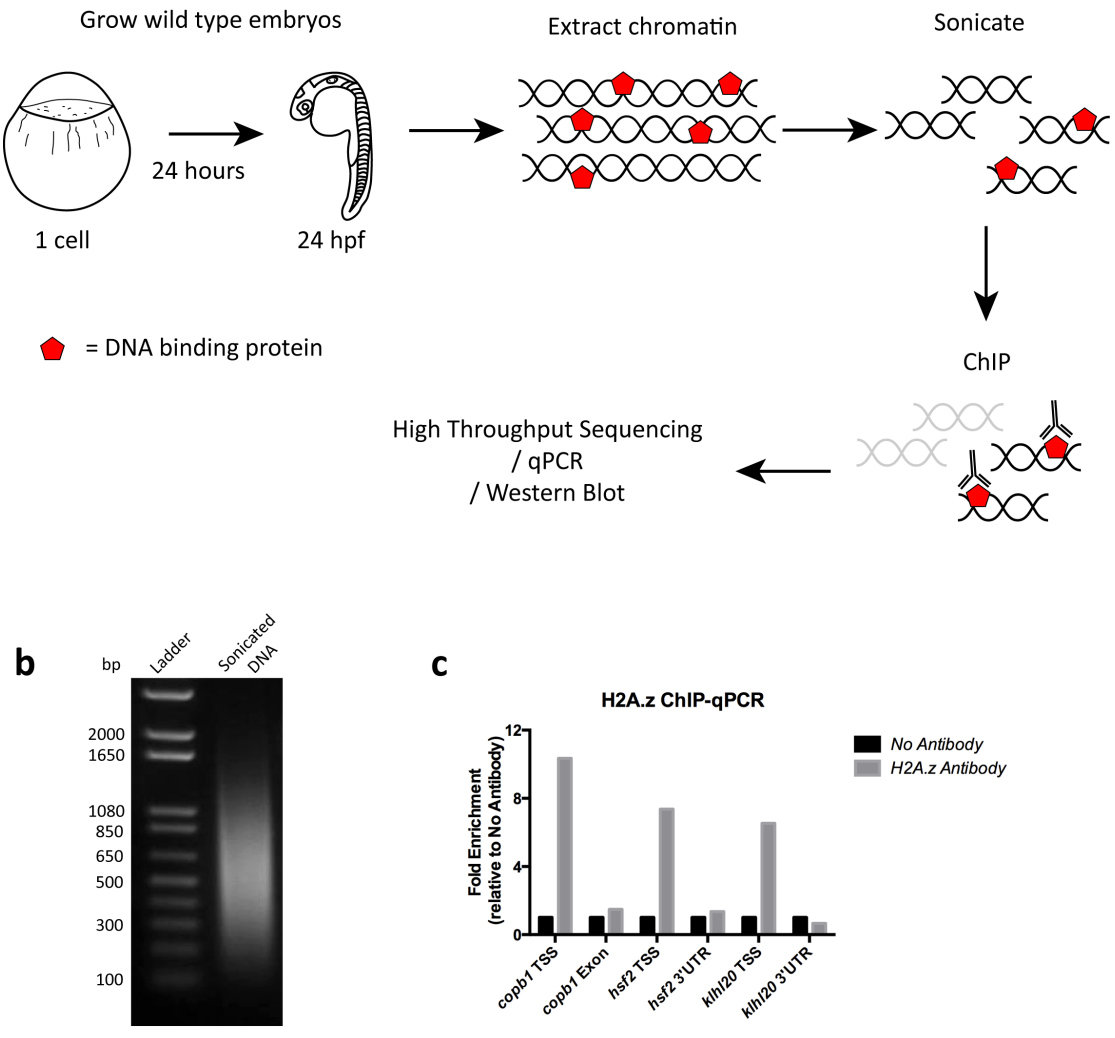
Figure 3-1: Testing the ChIP protocol for H2A.z

a: An overview of the ChIP procedure. 500 wild type embryos are grown to 24 hpf and the whole embryos are cross-linked to link the DNA-interacting proteins to the DNA. The embryos are lysed and chromatin extracted as a pool. The DNA is sheared by sonication into a suitable fragment size distribution (200 – 1000 bp for ChIP-PCR, 0 – 400 bp for ChIP-Seq). ChIP is carried out by the addition of an antibody specific to the protein of interest to 10 µg sheered chromatin. DNA fragments that are bound by this protein are isolated by immunoprecipitation, reverse cross-linked, purified and subject to analysis. This analysis can be high throughput sequencing in the case of ChIP-Seq, by qPCR for ChIP-PCR or by a Western Blot.

b: Chromatin was sonicated to produce fragments between 200 – 1000 bp. Sonication was carried out using a Diagenode Bioruptor for a total of 10 minutes on high setting using 30 second intervals. After reverse cross-linking, 1 µg of the sonicated DNA was run on a 1% agarose gel.

c: ChIP-PCR for H2A.Z shows enrichment at the TSS for examined genes. Fold enrichment relative to the no antibody control is shown.

Figure 3-1


a
Grow wild type embryos
24 hours
1 cell
24 hpf
Extract chromatin
Sonicate
= DNA binding protein
ChIP
High Throughput Sequencing
/ qPCR
/ Western Blot
b
bp
Ladder
Sonicated DNA
2000
1650
1080
850
650
500
300
100
c
H2A.z ChIP-qPCR
Fold Enrichment (relative to No Antibody)
12
8
4
0
No Antibody
H2A.z Antibody
copb1 TSS
copb1 Exon
hsf2 TSS
hsf2 3'UTR
klhl20 TSS
klhl20 3'UTR

that the antibody is capable of detecting both proteins. To confirm this, I transfected cultured cells with epitope tagged versions of Plzfa and Plzfb and carried out immunohistochemistry (IHC) using the zfPlzf Ab (Figure 3-2). Both Plzfa and Plzfb were found expressed in a speckled nuclear pattern (Figure 3-2h' & l') as previously reported for mouse and human Plzf (Reid et al., 1995, Bernardo et al., 2007).

ChIP against endogenous Plzf

As described later in Chapter 4.2, the zfPlzf Ab functions well in IHC using paraformaldehyde fixed zebrafish embryos, through this does not guarantee that it is ChIP-grade. Previous studies have identified ChIP-grade commercially available Plzf antibodies for use in human cells (Doulatov et al., 2009, Rice et al., 2009) and I therefore wanted to also determine whether they would be suitable for zebrafish ChIP. Information about these antibodies is provided in Figure 3-3a. Despite the relatively high sequence similarity between the epitope of the commercial antibodies, designated mPlzf and hPlzf, and the zebrafish Plzfa and Plzfb protein, I was unsuccessful in my attempts to use them to detect endogenous protein by IHC in zebrafish (*data not shown*).

In order to test these antibodies for ChIP I aimed to perform a similar ChIP-PCR experiment as I used for H2A.Z and confirm locus specific enrichment at sites bound by Plzf. In planning this, however, I encountered a second difficulty. As far as I am aware, there are no known targets directly bound by Plzf in the zebrafish genome and therefore I did not have any positive control regions with which I could test the antibodies. Instead, I turned to known Plzf targets found in mouse and human studies and compared them to the homologous regions within the zebrafish genome. Of those studied, *crabp1*, *hoxD11* and *ccna2* regulatory elements contain Plzf binding sites that are highly conserved in zebrafish (Figure 3-3b).

The combined evidence that these regions are well-characterised targets of Plzf in other organisms and are conserved in zebrafish meant that I used these as my positive control regions for ChIP-PCR. For negative control regions, I selected loci within the exons of three different genes that Plzf is not believed to bind to (*deltaA*, *ephrinA1*, *elavl3*). ChIP-PCR data showing the fold enrichment obtained from using the three described antibodies relative to the no antibody control at the positive and

Figure 3-2: Characterising the zfPlzf Ab

Cartoon showing that the zfPlzf Ab was raised against the Plzfa BTB domain. 112/121 of the amino acids between the two BTB domains are identical.

a-l: HEK 293 cells were transfected with either an Myc tagged control plasmid (a-d), HA tagged Plzfa (e-h) or Myc tagged Plzfb (i-l) and stained with antibodies against their respective tags. The zfPlzf Ab recognised both Plzfa and Plzfb, which are found in a nuclear speckled pattern (arrows in h' and l').
Scale bar = 5 μm.

Figure 3-2

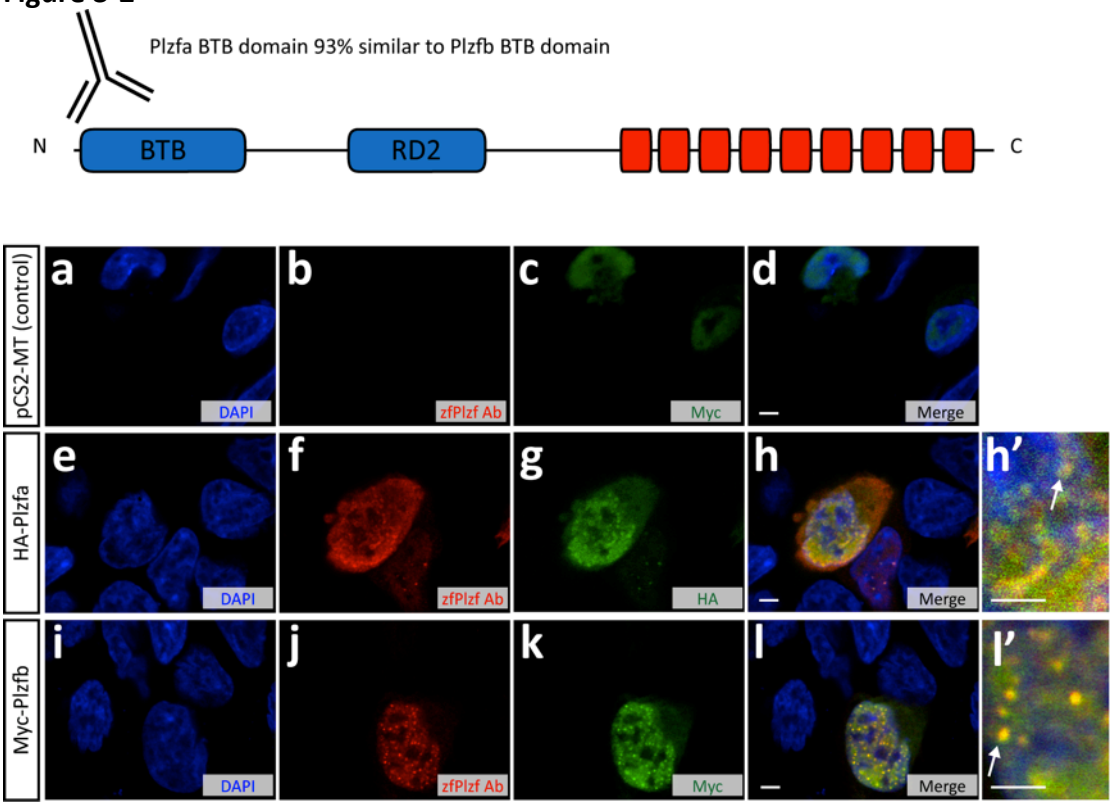

Plzfa BTB domain 93% similar to Plzfb BTB domain
N
BTB
RD2
C
pCS2-MT (control)
HA-Plzfa
Myc-Plzfb
a
b
c
d
e
f
g
h
h'
i
j
k
l
l'
DAPI
zfPlzf Ab
Myc
HA
Merge

Figure 3-3: ChIP-PCR attempts for endogenous Plzf

a: Relevant details of the three antibodies used for Plzf ChIP-PCR. Similarity to the zebrafish Plzfa and Plzfb proteins are shown as percentages of amino acid identity. Evidence that the mPlzf and hPlzf antibodies are ChIP Grade is provided by Doulatov et al. (2009) and Rice et al. (2009), respectively.

b: Homology between Plzf target regions within human, mouse and zebrafish genomes for Plzf target genes. The Plzf consensus binding sequence shown is the one defined by Li et al. (1997). The Plzf binding site within *crabp1*, *ccna2* and *hoxD11* regulatory sequences were described by Guidez et al. (2007), Yeyati et al. (1999) and Barna et al. (2002) respectively.

c-e: Plzf ChIP PCR results for potential positive control regions using the three candidate antibodies. Fold enrichment relative to the no antibody control is shown. No enrichment is observed for any of the loci analysed with any of the antibodies.

Figure 3-3

a

Name	Epitope	Similarity to Plzfa/Plzfb	ChIP Grade?
zfPlzf	zebrafish Plzfa BTB domain	100% / 96%	Unknown
mPlzf	Full length mouse Plzf	79% / 71%	Yes
hPlzf	N-terminal human Plzf	74% / 61%	Yes

b

Plzf consensus binding site : A T/G G/C T A/C A/C A G T

crabp1

```
Human      ggttggaaatgcggcgattt-cttctCATGTCATGa-------agctattt
Mouse      agttagaagtgtggggattttcttccCATGTCACAa-------agctattt
Zebrafish tgtcagaagtgccaaaaattgcaaatCGTATCATCatttatttggattatt
            **  *** **     * ** *    * * ***  *         *  **
                                     Plzf BS
```

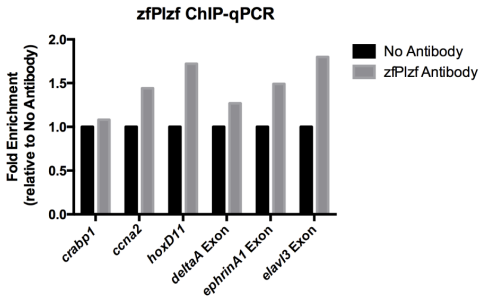
ccna2

```
Human      gccttgaatgACGTCAAGgccgcgagcgcttt
Mouse      gcttctggtgACGTCACGgactccgacgccct
Zebrafish acccctgatgACGTTGCCagtagccgcattct
           *      ******           *    *
                  Plzf BS
```

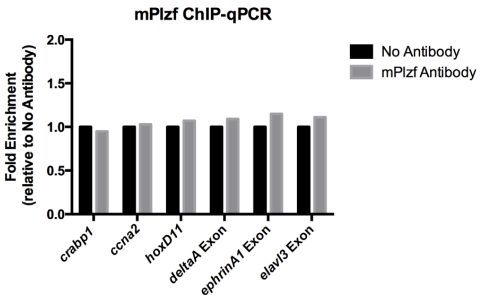
hoxD11

```
Human      tgcctttgaacttccaaaATGTCAAGgtcatcacctttaacctttctgaataat
Mouse      tgcctttgaacttccaaaATGTCAAGgtcatcacctttaacctctctgaataat
Zebrafish tgtatttgaacttctgtaATGTCAAGgtcgtcacccttaacctttttgaataag
           **  **********   ************ ***** ******* * *******
                           Plzf BS
```

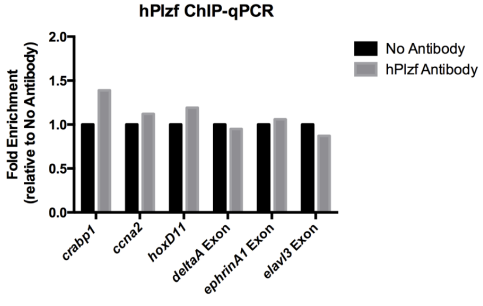
c



d



e



negative control regions is displayed within Figure 3-3c-e. The analysis did not show any enrichment at the positive control regions to suggest that any of the antibodies are working for ChIP in zebrafish. I cannot at this stage distinguish whether this is due to the antibodies not being ChIP-grade, whether the protocol requires optimisation for the antibodies to detect Plzf, or if it is because Plzf does not bind the selected regions in the zebrafish genome.

ChIP using tagged Plzf

The difficulty in performing ChIP-Seq for endogenous Plzf in zebrafish is that I lack a proven ChIP-grade antibody and cannot easily test the antibody due to the absence of positive control regions. In order to remove one of these caveats, I decided to inject an epitope tagged version of Plzfa into the zebrafish and then use a commercial, ChIP-grade antibody to detect the tag. Analysis of the tagged Plzf-bound regions by ChIP-Seq would provide me with a list of positive control regions which I can then use to re-evaluate the use of the antibodies shown in Figure 3-3a for ChIP. In theory, overexpressing tagged versions of transcription factors increases the possibility of false positives being identified where the excess protein leads to ectopic binding within the genome (Kidder et al., 2011). In practice, there is at least one example where comparison ChIP experiments between the overexpressed and endogenous transcription factor didn't reveal any increases in non-specific binding (Yao et al., 2013). Furthermore, the lack of antibodies available for zebrafish has led to other groups carry out ChIP-Seq on zebrafish embryos overexpressing tagged transcription factors (Xu et al., 2012). This method therefore seemed suitable for detecting Plzf bound target DNA.

In order to find a tagged version of Plzf that is suitable for ChIP experiments, I tested various different constructs. A N-terminal hemagglutinin (HA) tagged version of Plzfa (HA-Plzfa) was used previously for overexpression experiments in zebrafish (Sobieszczuk et al., 2010). A C-terminal FLAG tagged construct (Plzfa-FLAG) had been generated within the lab for structural and functional studies (Mohammed Ismail, unpublished). In addition, inspired by the successful ChIP-Seq experiments performed by Xu et al. (2012), I built a C-terminal Myc tagged construct (Plzfa-Myc). I aimed to test each construct for protein stability, whether they could be detected

after ChIP by Western Blot and whether they are able to induce the Plzfa overexpression phenotype.

I generated RNA for each of the constructs and injected them individually into wild type embryos. In order to confirm that the tagged protein is correctly translated and that the tags can be detected by IHC, I analysed expression within embryos at 3.5 hpf. At this pre-gastrulation time point, the exogenous overexpressed RNA has started to be translated and there is no endogenous Plzfa expression (Figure 3-4a-d). All of the three constructs tested showed nuclear Plzfa expression revealed by IHC with the zfPlzf Ab (Figure 3-4f, j & n). Both the Myc and HA tag were clearly detectable by IHC (Figure 3-4g & o), whilst the FLAG tag was undetectable. Therefore I decided to go ahead with just the HA-Plzfa and Plzfa-Myc constructs.

Next, I wanted to determine whether the tagged proteins were still detectable by IHC at a stage relevant to Plzf function. Whilst previously I had optimised the ChIP technique for 24 hpf embryos, now that I was injecting a specific amount of RNA into the one cell stage embryos, I was concerned that levels of protein expression would be relatively low by 24 hpf. Therefore, I decided to carry out experiments with the 14 hpf embryos, where previous experiments have shown endogenous Plzfa to have a role in regulating neurogenesis and when there is an overexpression phenotype (Sobieszczuk et al., 2010).

At this stage in development, I could observe exogenous nuclear Plzfa protein in embryos injected with RNA for *HA-plzfa* or *plzfa-Myc*. (Figure 3-4v & z). However, whilst I was able to detect the presence of the Myc tag, I couldn't detect the HA tag (Figure 3-4w & aa). I know that the HA tag can be observed by IHC (Figure 3-4g) and therefore this result suggests that the epitope is not accessible to the antibody at this stage in development. This result was confirmed by taking zebrafish lysates containing overexpressed proteins and performing Western Blots with the appropriate antibodies (*data not shown*). I was unable to reliably detect the HA tag within the HA-plzfa injected embryo lysate, further suggesting that this construct isn't suitable for detection of the epitope.

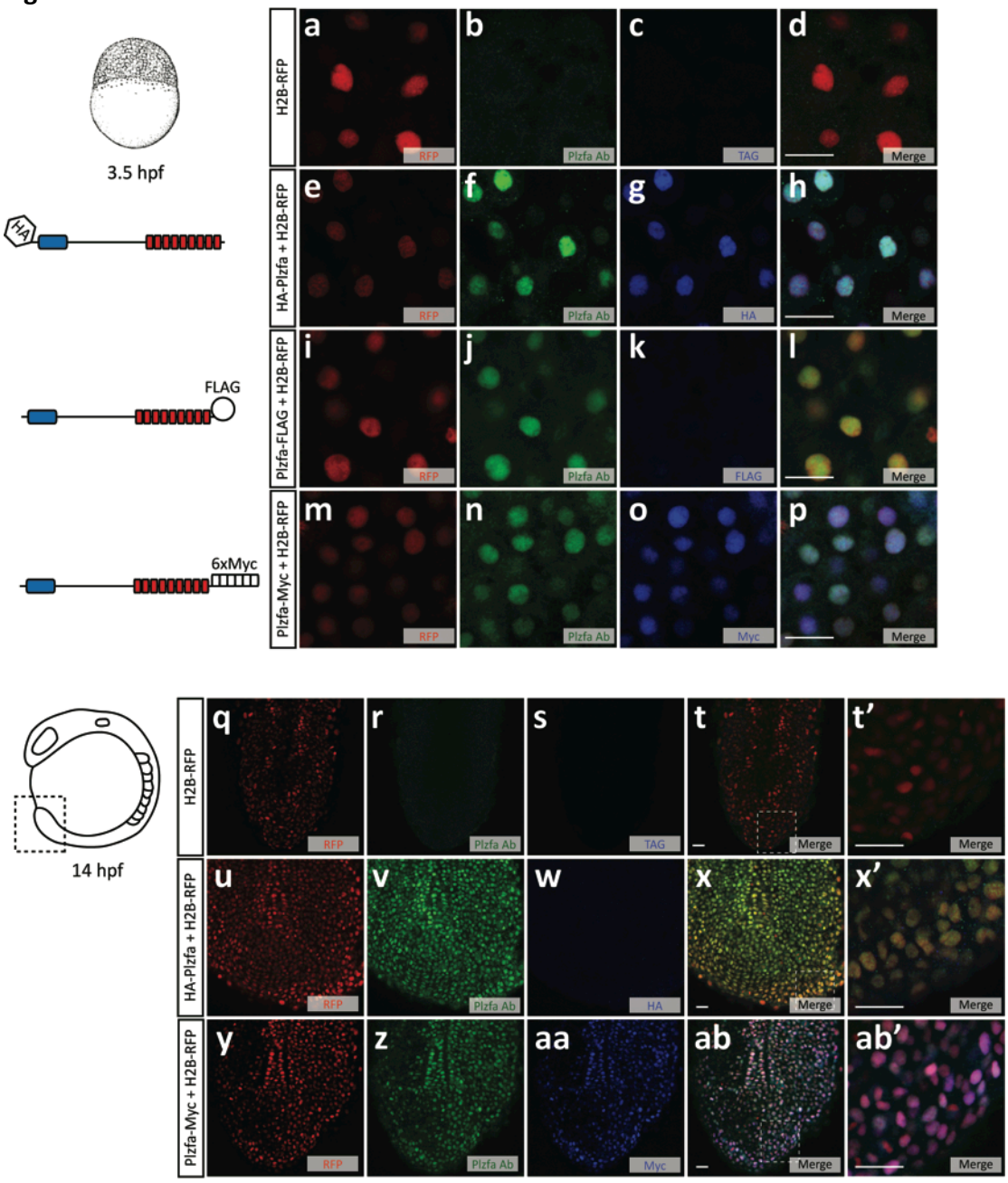
Figure 3-4: Expression of tagged Plzf constructs in zebrafish embryos

a-p: HA-Plzfa and Plzfa-Myc overexpressed protein is detectable by IHC in zebrafish. Wild type embryos were injected with 50 pg of one of the three tagged Plzfa constructs shown and analysed by IHC at 3.5 hpf. The zfPlzf Ab recognised all three constructs (f,j & n) and the HA and Myc antibodies recognised their respective tagged constructs (g & o). The FLAG antibody was unable to recognise overexpressed Plzfa-FLAG (k) and therefore this construct was not taken forward. No endogenous Plzfa protein was detected at this stage in the control injected embryos (b).

q-ab: Plzfa-Myc is detectable at within the 14 hpf zebrafish. Embryos injected with the HA-Plzfa and Plzfa-Myc constructs were developed until 14 hpf and expression was analysed at the tailbud. There is no endogenous Plzfa protein in this location in the control injected embryos (r). Overexpression of either construct leads to nuclear Plzf expression (v & z) but only the Myc tag is detectable by IHC (w & aa). Zoomed images of the boxed regions are shown in t',x' and ab'.

Scale bar = 20 μm.

Figure 3-4


3.5 hpf
HA
FLAG
6xMyc
H2B-RFP
HA-Plzfa + H2B-RFP
Plzfa-FLAG + H2B-RFP
Plzfa-Myc + H2B-RFP
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
RFP
Plzfa Ab
TAG
HA
FLAG
Myc
Merge
14 hpf
q
r
s
t
t'
u
v
w
x
x'
y
z
aa
ab
ab'

The analysis so far suggests that the Plzfa-Myc protein is stable at 14 hpf and therefore is a good candidate for performing ChIP-Seq. I aimed to confirm that I could immunoprecipitate the tagged protein from chromatin isolated from injected embryos. Before doing this, I had to optimise the sonication conditions required for fragmenting DNA to a range suitable for ChIP-Seq (0-400 bp). I sonicated the isolated chromatin from injected embryos at 14 hpf for increasingly long times (Figure 3-5a). I found that I could obtain ideal conditions when sonication was performed for 40 minutes total, and that I could reliably reproduce this result with multiple biological replicates (Figure 3-5b). This reproducibility is important because, in order to generate the required amount of starting material for ChIP-Seq, chromatin was pooled from multiple injections performed on different days.

With the conditions optimised, I carried out ChIP using the ChIP-grade antibody against the Myc epitope used previously in zebrafish (Xu et al., 2012). Still lacking positive regions that Plzf binds within the zebrafish genome, I instead turned to using ChIP followed by a Western Blot in order to confirm the antibody can pull down Plzfa-Myc. After performing ChIP, I extracted the protein and used a second anti-Myc antibody in order to see whether I could detect the presence of the appropriately size Plzfa-Myc protein (86.6 kDa). Figure 3-5c shows that Plzfa-Myc was evident within the ChIP sample that had been obtained from embryos injected with RNA encoding the construct, although at a slightly higher molecular weight than found in the input lanes. Reassuringly, no Myc tag was detected within the uninjected samples and the nuclear protein HDAC1 was not present within the ChIP sample.

I used the same method to determine whether the zfPlzf Ab is suitable for ChIP. I carried out ChIP with the zfPlzf Ab on embryos injected with Plzfa-Myc and used the anti-Myc antibody to detect whether the protein is present in the sample (Figure 3-5d). Whilst I was able to observe the tagged protein within the lysate and chromatin input, it was not detected in the ChIP sample. This shows that, under the conditions used within my ChIP-Seq protocol, the zfPlzf Ab is not suitable for ChIP experiments.

Figure 3-5: Plzfa-Myc suitability for ChIP in zebrafish

a: Optimising sonication conditions for ChIP-Seq. Pools of 100 12 hpf injected embryos were sonicated for different lengths of time (0 – 40 minutes) as specified. After sonication, the DNA was reverse cross linked and purified. The range of DNA fragments becomes more compact as sonication length increased and 40 minutes sonication time was ideal for ChIP-Seq experiments (0-400 bp).

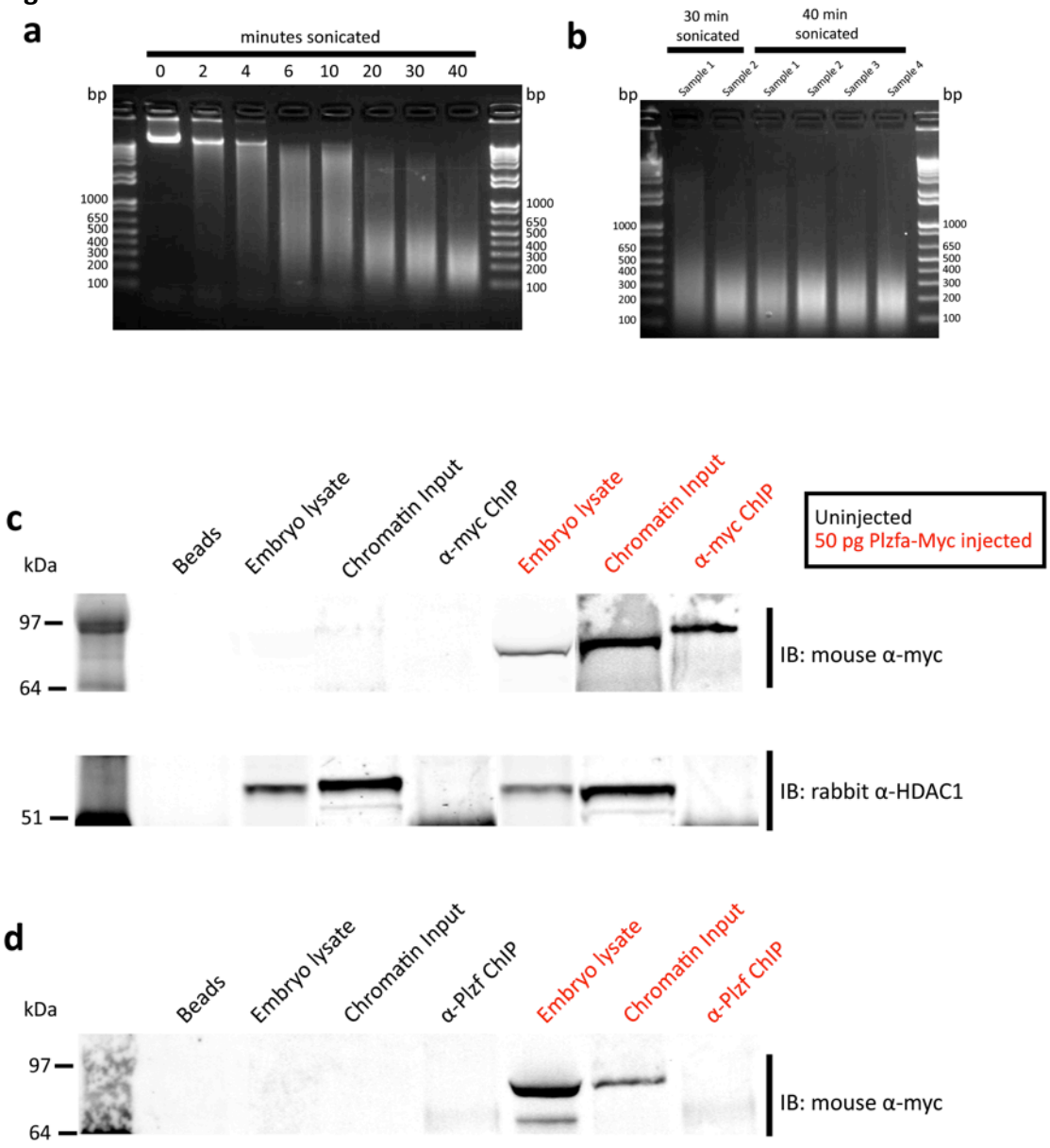
b: Optimal sonication conditions are highly reproducible. Each sample is a different biological replicate sonicated for either 30 or 40 minutes as indicated. For both gels, 1 µg of DNA was run on a 1.5% agarose gel.

c: Plzfa-Myc is detectable in a Western Blot after ChIP using a Myc antibody. Total protein was extracted from samples that had gone through the IP process but contained no chromatin ('beads', 20 µl), whole embryo lysates (30 µg), isolated chromatin taken before the IP ('chromatin input', 30 µg) or the total ChIP sample (α-myc ChIP, 20 µl). Samples in black were obtained from uninjected embryos and samples in red were obtained from embryos injected with 50 pg *plzfa-myc* RNA. The immunoblot (IB) against HDAC1 is shown as a control for the presence of a nuclear protein within the embryo lysates and chromatin input, but not present in the ChIP sample. For an unknown reason, the size of the Plzfa-Myc protein appears slightly higher within the ChIP sample compared to same protein in the lysate and chromatin input.

d: Plzfa-Myc is undetectable in a Western Blot after ChIP using zfPlzf Ab. Samples were run as above, but the ChIP was carried out with the zfPlzf Ab. No band is observed at the expected size to indicate Plzfa-Myc is detected in the ChIP sample.

Expected size of Plzfa-Myc is 86.6 kDa, HDAC1 is 56 kDa and the IgG heavy and light chains are 55 and 25 kDa respectively.

Figure 3-5


a
minutes sonicated
0 2 4 6 10 20 30 40
bp
1000
650
500
400
300
200
100
b
30 min sonicated
40 min sonicated
Sample 1
Sample 2
Sample 1
Sample 2
Sample 3
Sample 4
c
kDa
Beads
Embryo lysate
Chromatin Input
α-myc ChIP
Embryo lysate
Chromatin Input
α-myc ChIP
Uninjected
50 pg Plzfa-Myc injected
97
64
IB: mouse α-myc
51
IB: rabbit α-HDAC1
d
kDa
Beads
Embryo lysate
Chromatin Input
α-Plzf ChIP
Embryo lysate
Chromatin Input
α-Plzf ChIP
97
64
IB: mouse α-myc

Testing functionality of Plzfa-Myc

Finally, I aimed to determine whether the Plzfa-Myc construct was functional. Previous overexpression experiments with Plzfa in zebrafish had shown a neurogenic phenotype whereby there is a decreased number of *neurog1* and *islet1* positive cells in *myc-plzfa* RNA injected embryos (Sobieszczuk et al., 2010) and a similar phenotype has been described in chicken embryos (Gaber et al., 2013). I planned to repeat these experiments in order to determine whether overexpression of the Plzfa-Myc construct would result in a similar phenotype. I aimed to use the previously published HA-Plzfa construct as a positive control for the phenotype.

I injected RNA coding either the HA-Plzfa or the Plzfa-Myc protein into 1 cell at the 2-cell stage zebrafish embryo along with a tracer to follow the RNA expression (Figure 3-6a). This was a similar experiment as to that described previously (Sobieszczuk et al., 2010) except that in the published work *lacZ* RNA was used as the tracer, whilst I injected RNA for the fluorescent protein H2B-Citrine. Injection of the RNA encoding either construct didn't lead to the expected decrease in the number of *islet1* expressing cells (Figure 3-6e-j).

The fact that neither the HA-Plzfa nor the Plzfa-Myc constructs were able to give a phenotype similar to that described was unexpected. One possible explanation for this result is that both constructs are functionally inactive within the zebrafish embryo. It is possible that the position of the Myc tag at the C-terminus, close to the zinc fingers of Plzf, may interfere with its DNA binding, whilst the HA tag at the N-terminus is close to the BTB domain and therefore may interfere with the activity of Plzf. There are also technical considerations that may be affecting the result, such as the difference in the method of detection described above and the amount of RNA injected.

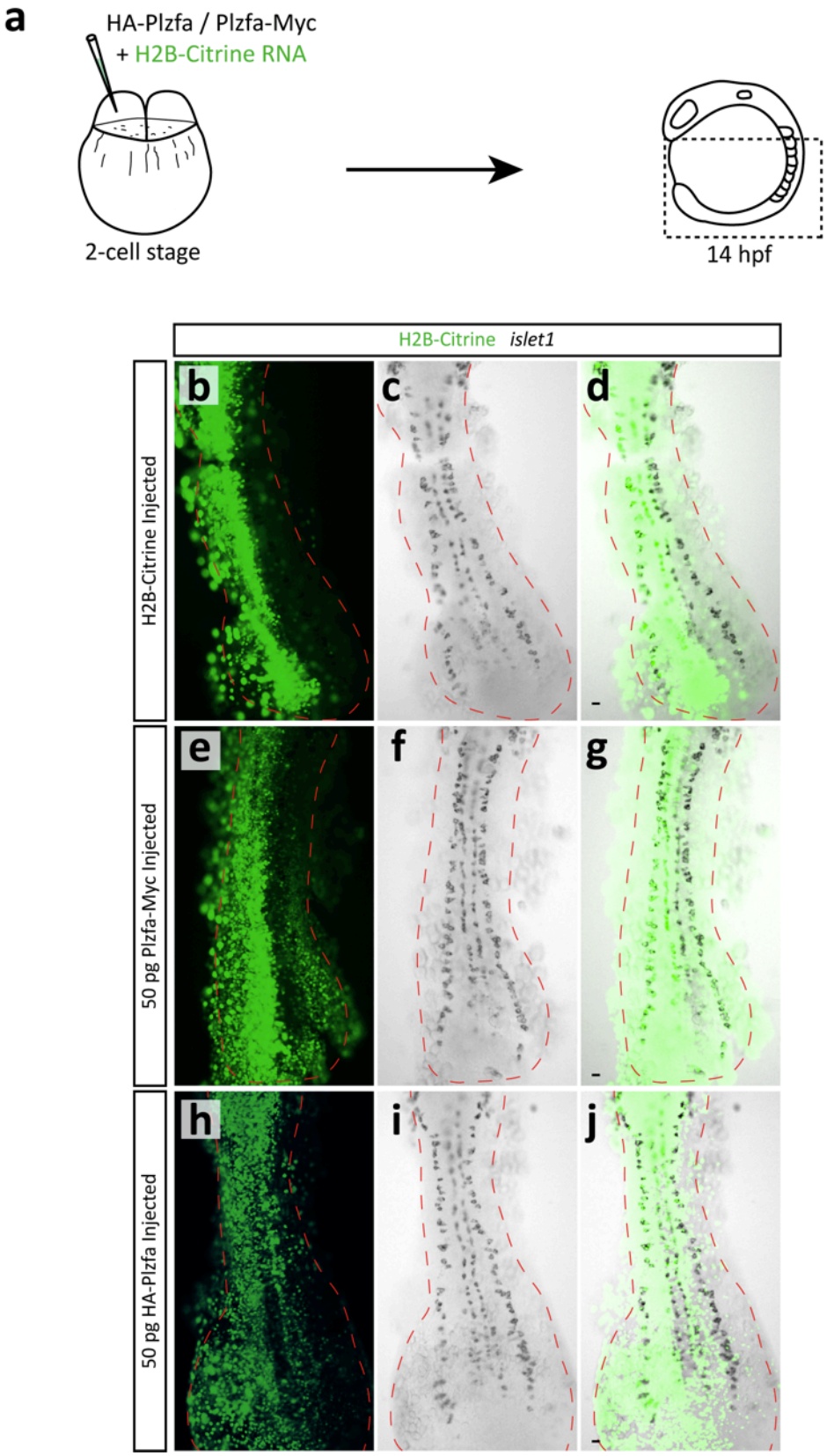
Figure 3-6: Testing functionality of constructs

a: 1 cell of a 2-cell stage embryo was injected with RNA encoding either HA-Plzfa or Plzfa-Myc. H2B-Citrine RNA was used as a tracer to follow the side of the embryo that contains the construct. Embryos were developed until 14 hpf when they were analysed by ISH for *islet1* expression.

b-j: 14 hpf embryos injected with either construct didn't display an overexpression phenotype. Control embryos (b-d) were injected with H2B-Citrine alone and the injected side didn't show any changes in *islet1* expression (0/10). Likewise, the Plzfa-Myc and HA-Plzfa RNA injected embryos did not show any changes in the number of *islet1* cells (0/23 and 0/24 respectively) when comparing the injected side of the embryo to the uninjected side. The dashed red line shows the rough outline of each embryo.

Scale bar = 20 μm.

Figure 3-6


a
HA-Plzfa / Plzfa-Myc
+ H2B-Citrine RNA
2-cell stage
14 hpf
H2B-Citrine islet1
H2B-Citrine Injected
50 pg Plzfa-Myc Injected
50 pg HA-Plzfa Injected
b
c
d
e
f
g
h
i
j

In the interest of being able to obtain a list of transcriptional targets for Plzf within the timescale of my PhD project, I made the decision to go ahead and carry out ChIP-Seq using the Plzfa-Myc injection experimental paradigm despite the lack of evidence that this construct is functional *in vivo*. Whilst waiting for results from the sequencing, I continued these functional experiments as described later.

3.3 ChIP-Seq using Plzfa-Myc

In order to generate enough starting material for Illumina sequencing I injected and extracted chromatin from 1200 embryos, producing around 85 µg of total DNA. 1% of the chromatin was taken as the input sample. The remaining material was split into 7 IP experiments using the Myc antibody and the resulting purified DNA was pooled together in 35 µl. The final concentration of DNA after ChIP was 354 pg/ µl and a total of 10 ng DNA from the ChIP sample and from the input sample was taken forward for library preparation by the NIMR HTS facility. The samples were run on the Illumina HiSeq 2000 platform.

Sequencing of the ChIP product generated 3.3 x 10^7 sequences and the input produced 1.5 x 10^7 sequences. These 40 bp sequences were found to be high quality (Figure 3-7a) and were mapped to zebrafish genome assembly (Zv9). The next step in the analysis is to use peak calling software to identify regions of ChIP enrichment, demonstrated in Figure 3-7b. I used model-based analysis of ChIP-Seq (Zhang et al., 2008b) on my data but the analysis failed, as it was unable to find the minimum number of pair peaks in order to build the model. I adjusted the available parameters but was still unable to successfully call peaks from the obtained sequencing data. This is described in more detail in the Discussion section below.

I turned to visualising the mapped reads from both the ChIP and input sample on a genome browser in order to gain an impression of any regions of enrichment. Scanning through the genome manually failed to identify any evidence of peaks. I selected some candidate regions where Plzf is known to bind in other genomes and saw no obvious enrichment within the ChIP dataset (Figure 3-7c). From my analysis, I conclude that the ChIP-Seq experiment failed to identify any regions that Plzf binds to within the zebrafish genome.

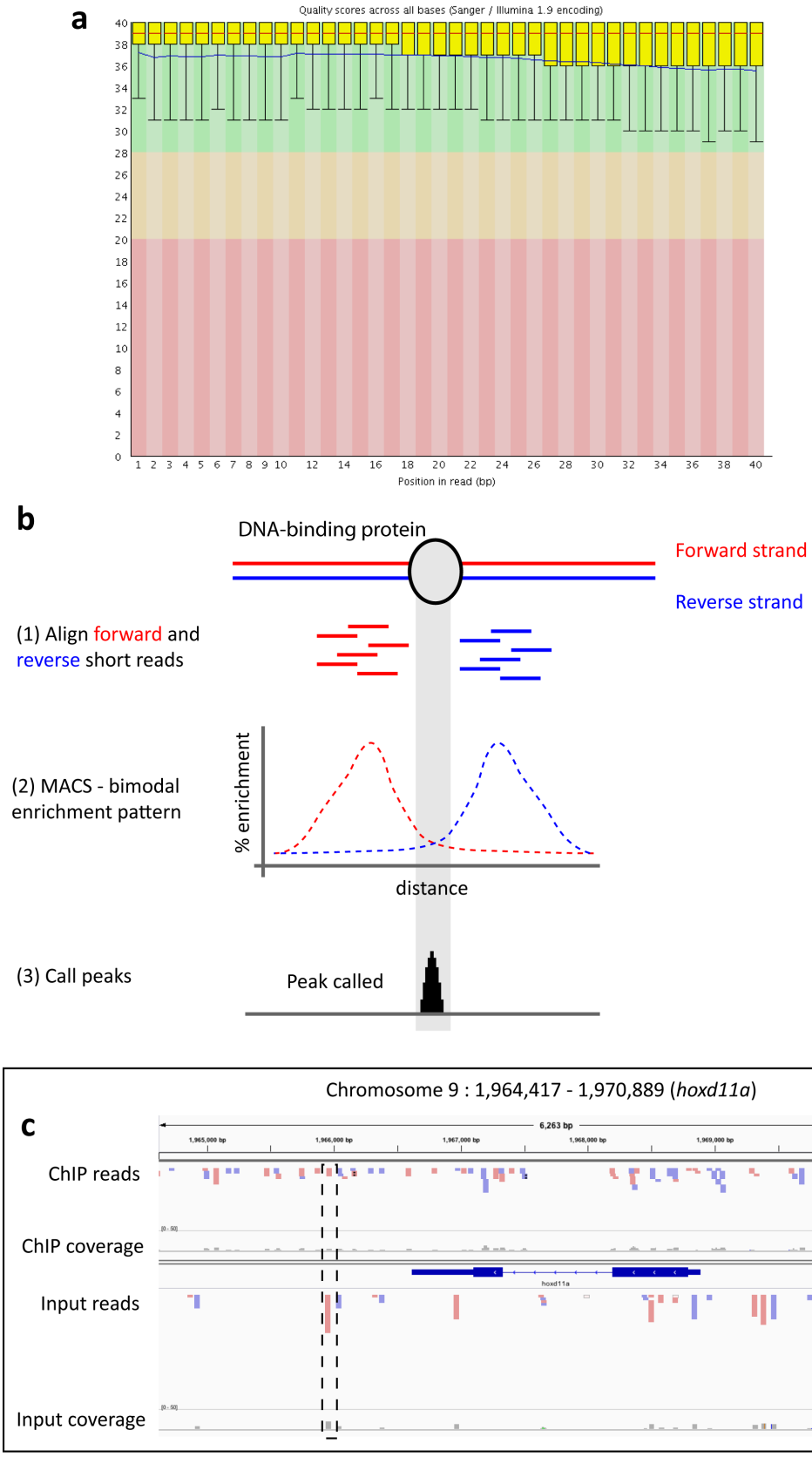
Figure 3-7: ChIP-Seq results for Plzfa-Myc

a: Per base sequence quality boxplot generated by FastQC utility. The yellow boxplots show base-calling quality scores (0-40) across all sequencing reads from the ChIP sample. Red line: median, box: interquartile ranges 25 – 75% and whisker: 10 – 90% percentile. Sequencing quality decreases as position increases, although overall the quality is very good, with a score > 30 meaning 99.9% confidence that the base is called correctly.

b: Cartoon showing a simplified demonstration of peak calling. A DNA-binding protein is shown to be interacting at a particular location in the genome. Sequencing results in a many short forward (red) and reverse (blue) tags that can be aligned (1) to a reference genome. Using the MACS software to call peaks (2), this initially determines the enrichment pattern of the forward and reverse reads at locations within the genome being mapped. MACS then uses this information to translate the enrichment of reads into a "peak" (3) and predict binding sites of the protein of interest.

c: Images taken from the IGV genome browser showing ChIP and Input reads mapped to the zebrafish genome at *hoxd11a*. 40 bp forward reads are shown in red and reverse reads in blue. Region boxed indicates potential Plzf binding site shown previously in Figure 3-3b. No enrichment of forward and reverse reads that would be indicative of a peak is observed at this location in the ChIP sample. Scanning through other regions of the zebrafish genome did not reveal any enrichment.

Figure 3-7


a
Quality scores across all bases (Sanger / Illumina 1.9 encoding)
Position in read (bp)
b
DNA-binding protein
Forward strand
Reverse strand
(1) Align forward and reverse short reads
(2) MACS - bimodal enrichment pattern
% enrichment
distance
(3) Call peaks
Peak called
c
Chromosome 9 : 1,964,417 - 1,970,889 (hoxd11a)
6,263 bp
ChIP reads
ChIP coverage
Input reads
Input coverage
hoxd11a

3.4 Troubleshooting the ChIP-Seq experiment

I aimed to determine the reason why I didn't obtain any Plzf binding regions from my data. As described, I was without evidence that the Plzfa-Myc construct was functional *in vivo*. This could mean that this protein is unable to bind the Plzf target regions within the zebrafish genome and therefore could account for the lack of useful data acquired from the sequencing. As previously described, the position of the tag might interfere with the ability of the constructs to bind DNA. The previously published work obtained the phenotype using an N-terminal Myc tagged Plzfa (Myc-Plzfa) (Sobieszczuk et al., 2010), and I wanted to confirm whether this protein is functional when RNA is injected into the zebrafish embryos.

I carried out a similar experiment to that shown in Figure 3-6, injecting Myc-Plzfa RNA such that it is expressed within half of the embryo and analysing at 14 hpf. I found that overexpression of Myc-Plzfa leads to a decrease in *neurog1* and *islet1* expressing cells (Figure 3-8a-i). This phenotype is reminiscent to the Plzfa overexpression phenotype described previously (Sobieszczuk et al., 2010). Previous analysis showed that the N-terminal HA tag of HA-Plzfa was undetectable by IHC at 14 hpf (Figure 3-4w) and therefore I wanted to check whether this was the case for the N-terminal Myc tag. I was able to reliably detect the presence of the Myc tag at 14 hpf in the injected embryos (Figure 3-8j-l). Overall, these results show that the Myc-Plzfa construct is stably expressed and functional *in vivo*.

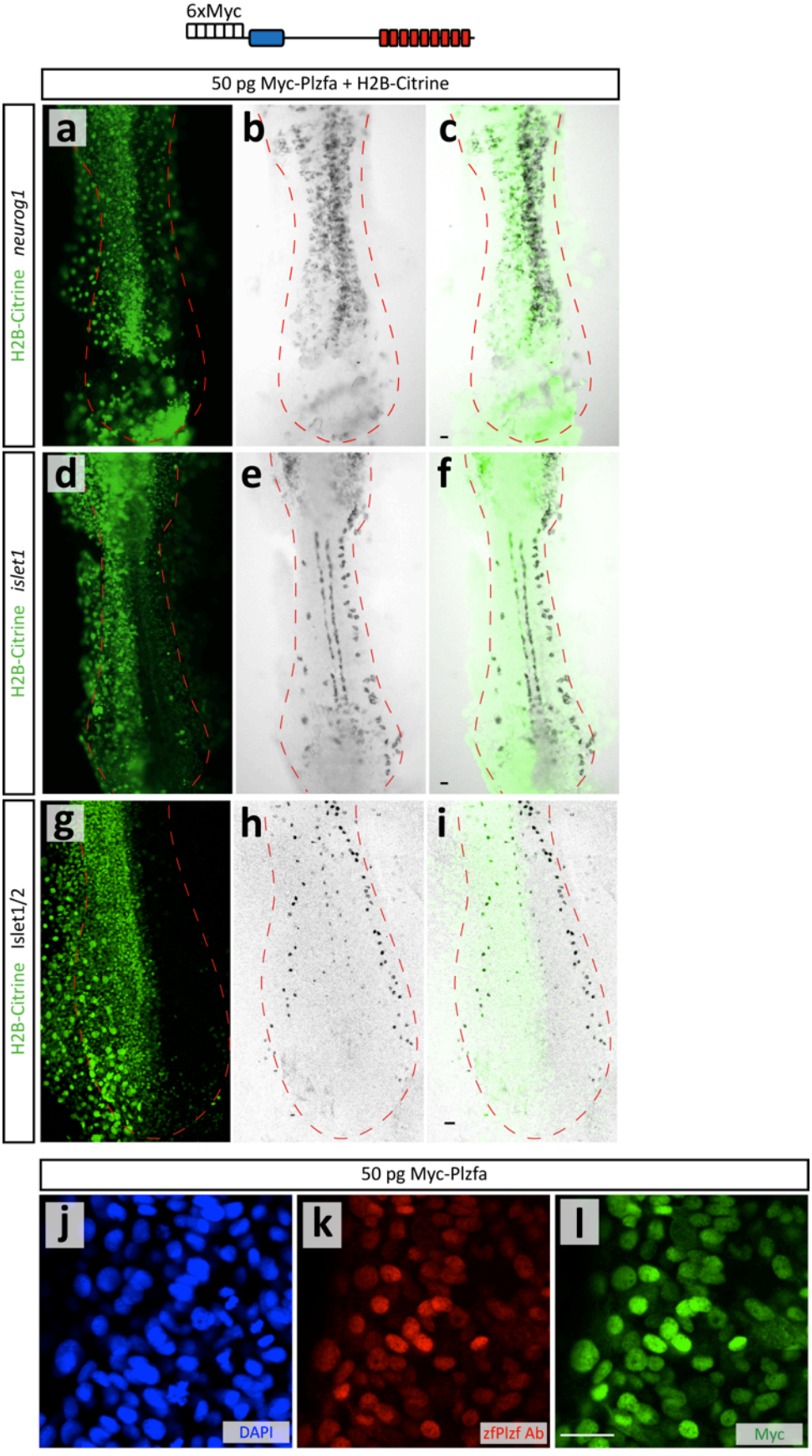
Figure 3-8: Myc-Plzfa overexpression causes a neurogenic phenotype

a-i: 1 cell of a 2-cell stage embryo was injected with RNA encoding Myc-Plzfa. Embryos were developed until 14 hpf when they were analysed by ISH for *neurog1* (a-c) and *islet1* (d-f), or by IHC for Islet1/2 protein expression (g-i). Overexpression of Myc-Plzfa lead to a down regulation of *neurog1* (20/35 embyros) and *islet1* (41/53 embryos) and Islet1/2 protein (6/6 embryos) in the injected side of the embryo. Images shown in g-i are maximum projections taken from confocal images. The dashed red line shows the rough outline of each embryo.

j-l: Confirmation that the Myc tag is detectable at 14 hpf. Confocal images shown are from an injected embryo, showing that the overexpressed protein is nuclear (j) and is recognisable by both the zfPlzf Ab (k) and the Myc antibody (l).

Scale bar = 20 μm.

Figure 3-8


6xMyc
50 pg Myc-Plzfa + H2B-Citrine
H2B-Citrine neurog1
H2B-Citrine islet1
H2B-Citrine Islet1/2
a
b
c
d
e
f
g
h
i
50 pg Myc-Plzfa
j
k
l
DAPI
zfPlzf Ab
Myc

3.5 Discussion

A significant amount of effort went into determining the ideal parameters for carrying out ChIP-Seq for Plzf in zebrafish embryos but overall I was unable to generate a list of transcriptional targets that are bound by Plzf within the timescale of PhD project.

Difficulties working with endogenous Plzf

Due to my inexperience with the ChIP technique, I initially aimed to test that I could perform the experiment using a known ChIP-grade antibody, H2A.Z, and known positive and negative control regions. Success with this experiment meant I was confident in my protocol and I proceeded to perform the same technique using the various Plzf antibodies described. The absence of positive regions that Plzf binds to within the genome meant that evaluating the antibodies for ChIP was difficult.

I attempted to identify conserved Plzf binding sites between species by DNA sequence comparison in order to obtain target loci where Plzf is likely to bind within the zebrafish genome. Evidence suggests that the majority of transcription factor binding events are species-specific despite conservation of gene function (Schmidt et al., 2010, Odom et al., 2007). Analysis by ChIP-PCR did not reveal any enrichment at the chosen loci when using any of the Plzf antibodies. The lack of enrichment could be due to the antibodies not functioning in ChIP, or alternatively it could be that Plzf does not bind the loci selected within the zebrafish genome.

Lessons learnt from epitope tagging

In order to circumvent the possibility that the antibody isn't working, I chose to evaluate the use of epitope tagged form of Plzf and make use of antibodies known to work in ChIP. A series of experiments were carried out to determine whether the various tagged constructs were detectable at an appropriate stage in zebrafish development, could be detected by ChIP and whether they were functional. From these experiments, it was found that the C-terminal tagged Plzfa-Myc protein was stably expressed and could be detected after ChIP experiments, but overexpression studies suggested that it wasn't functional. Nevertheless, the fact that this epitope tagged protein was detectable and expressed within the nucleus meant that I

elected to go ahead and carry out the ChIP-Seq experiment despite the lack of functional data.

The purpose of the ChIP-Seq experiment is to obtain sequenced reads that can be aligned to a genome and used to identify where a protein binds. Reads that directly relate to locations where the protein binds have to be separated from 'background reads' that are sequenced from DNA that is not specific to the protein-binding site. In order to do this, computation methods identify regions of read enrichment, which are subsequently translated into peaks. It is important to also carry out sequencing of the sonicated DNA sample that has not undergone the ChIP process with a specific antibody, known as the input sample. The resulting reads from the input sample should only correspond to the background reads and therefore can be directly compared with and subtracted from the reads from the ChIP sample.

I chose to use the publically-available peak calling software MACS (Zhang et al., 2008b). Since ChIP-DNA fragments are equally likely to be sequenced from both ends, the tag density should produce a bimodal enrichment pattern (illustrated in Figure 3-7b). The spatial distribution of these sequenced forward and reverse tags correlates with a protein DNA-binding position. By inputting parameters such as the length of the reads and size of the sonicated fragments, MACS identifies these spatial enrichment patterns and builds a model with this information in order to call peaks and therefore locate binding sites.

When I inputted my data into the MACS software, it was unable to identify a sufficient number of these spatially distributed enrichment patterns in order to build the model and subsequently call peaks. The most likely reason for this is that the DNA immunopecipitated and sequenced in my epitope tagged ChIP-Seq experiment does not correspond to the Plzf binding site, and instead consists of the aforementioned background reads. Although it was not feasible to manually inspect the entire zebrafish genome to see if I could identify locations where the reads were enriched, observing the mapped reads at potential Plzf binding sites (such as those shown in Figure 3-3b) didn't reveal any indication of enrichment, or any obvious differences with the mapped reads obtained from the input sample. The evidence

therefore suggests that the ChIP-Seq experiment failed to identify genomic loci where the tagged-Plzfa protein binds.

A potential reason for this failure is that the C-terminal Plzfa-Myc construct is unable to bind to the target loci. Evidence that this comes from the fact that I didn't observe the previously described overexpression phenotype (Sobieszcuk et al., 2010) when I injected the RNA encoding this tagged protein. I therefore carried out preliminary functional and expression experiments using an N-terminal tagged Myc-Plzfa construct. Using this construct resulted in evidence that the N-terminal tagged protein was functional and that the tag was detected by immunohistochemistry. Future experiments would be to go ahead and carry out ChIP-Seq using this epitope tagged protein in order to identify transcriptional targets of the Plzfa protein.

With the benefit of hindsight, I should have evaluated both the N-terminal and C-terminal Myc tagged constructs to determine which to take forward for ChIP-Seq. Based upon the known domains of the Plzf protein, it is possible that adding the six Myc epitopes close to the zinc finger motifs is disrupting the DNA binding activity of Plzf. When designing of the Plzfa-Myc construct, I took into account the difficulties in detecting the N-terminal HA tag described in the text and also that previous publications have demonstrated that adding a HA tag to the C-terminus of Plzf doesn't disrupt its function (Seidel et al., 2010, Zaade et al., 2013).

A further aspect to consider is that the tagged protein has to compete with the wild-type Plzfa protein that normally occupies the Plzf binding sites in the zebrafish genome. The C-terminal Plzfa-Myc protein may therefore be unable to outcompete the wild-type protein and this could be why it is unable to bind the targets in the genome. This concern could be alleviated by carrying out the same ChIP-Seq experiment in embryos that lack the wild-type Plzfa protein, such as those generated in Chapter 5.

Future considerations

The decision to use whole zebrafish embryos at the source of cells for the ChIP experiment was largely made because my functional experiments were being carried out in the same system. Other possibilities include working with embryonic or neural stem cells such as those where Plzf is highly expressed (Elkabetz et al.,

2008). However, cultured cell lines are not available for zebrafish and therefore this would entail determining transcriptional targets within a different species, which would not necessarily be targets within the zebrafish. Furthermore, the long-term maintenance of cells in culture can deregulate their spatial and temporal identity and therefore performing ChIP-Seq using this system may not provide the full spectrum of transcriptional targets across the developing embryo (Conti and Cattaneo, 2010).

A potential caveat to performing the ChIP-Seq experiments using whole zebrafish embryos is that Plzf is likely to be regulating gene expression in tissues other than the nervous system. In order to confirm that any identified Plzf targets are being regulated within the neural progenitor cells, the spatial expression of the targets would need to be analysed in embryos where Plzf function has been manipulated. Combing the ChIP-Seq with fluorescence-activated cell sorting allows the tissue specific identification of targets (Bonn et al., 2012). Therefore, using transgenic zebrafish lines such as the *gfap:gfp* line, used in Chapter 4, could help obtain neural specific targets .

The described experiments makes the assumption that Plzfa and Plzfb proteins bind to the same region in the zebrafish genome, based on their functional redundancy proposed later in Chapter 4.4. Overexpressing a tagged version of Plzfb could be used in order to determine the extent of overlap of targets between the two paralogues. As the ChIP-Seq experiment failed to generate data, I am unable to comment on whether the choice to overexpress Plzf leads to increased background signal. The recent advent of genome editing technology mean that, using techniques such as described within Chapter 6, it would be possible to add epitope tags to endogenous genes such as has been done within the mouse (Zhang et al., 2008a).

Conclusions

Knowledge of the transcriptional targets that are regulated by Plzf within neural progenitors would significantly improve our understanding of how Plzf carries out its functions. Whilst I was unsuccessful in my attempts to achieve this aim, a considerable amount of work has gone into determining the various parameters

required to carry out ChIP-Seq in zebrafish. This work will therefore be of value to future experiments that aim to characterise genes regulated by Plzf.

4 Investigating the role of Plzf in neural progenitor maintenance

4.1 Introduction

The Plzf transcription factor is known to regulate progenitor maintenance in several tissues (Suliman et al., 2012). The expression pattern for Plzf in the vertebrate nervous system was first described almost 20 years ago (Cook et al., 1995) and yet it is only recently that a functional role in this tissue has begun to be elucidated. Previous work within the Wilkinson lab identified Plzfa as a widely expressed inhibitor of the initial wave of neurogenesis in zebrafish, but suggested that it is entirely redundant to Notch-mediated lateral inhibition (Sobieszczuk et al., 2010). Whilst my PhD project was underway, another group showed a function for Plzf in the maintenance of neural progenitors within the chicken and mouse spinal cord by regulating the cells responsiveness to FGFs (Gaber et al., 2013).

I was interested in whether there are proteins other than Notch that could be functionally redundant to Plzfa in the zebrafish. The highly similar RNA expression patterns for *plzfa* and its paralogue, *plzfb*, described previously (Sobieszczuk et al., 2010), made Plzfb an interesting candidate and I aimed to explore its expression and function further. The previous results in zebrafish focused on early stages of neurogenesis and I was interested in whether there is also a role at later stages during embryogenesis.

In this chapter I shall firstly describe the RNA and protein expression patterns observed for *plzfa* and *plzfb* during zebrafish development, focusing on its expression within the nervous system. Choosing the 44 hpf hindbrain as a model for investigating neurogenesis, I shall explore in more detail where Plzf is expressed during the neurogenic cascade, comparing it with the expression of markers of neural progenitors, differentiating cells and differentiated neurons. Using morpholino knockdown techniques, I describe a phenotype relating to a defect in progenitor maintenance observed when the functions of both *plzfa* and *plzfb* are disrupted. Finally, I use these results to revisit the feedback loop involving Plzf and Btbd6a previously proposed (Sobieszczuk et al., 2010).

4.2 Examining expression of the *plzfa* and *plzfb* paralogues

I began by examining the similarity between the *plzfa* and *plzfb* genes. *plzfa* is found on chromosome 21 within the zebrafish genome and its paralogue exists on chromosome 15. Within both the mouse and human genomes, the *Plzf* gene is composed of 7 exons and encodes a protein with 9 zinc fingers. Examination of the zebrafish genome (Zv9 assembly) reveals that *plzfb* has the same genomic structure as the human and mouse orthologue whilst only 5 exons for *plzfa* can be identified (Figure 4-1). The 2 missing exons in the *plzfa* gene encode zinc fingers 4 – 7 of the Plzfa protein. Evidence that the Plzfa protein has the expected 9 zinc fingers comes from the existence of full-length *plzfa* cDNA clones derived from RNA extracted from wild type zebrafish. (IMAGE ID: 3815539 & 100060713). My interpretation of these observations is that the *plzfa* gene is composed of 7 exons as seen in other species and is incorrectly annotated in the current zebrafish assembly.

The *plzfa* open reading frame contains 2016 bp, which encode 671 amino acids, similar to the mouse protein at 673 amino acids. The Plzfb protein is slightly shorter at 659 amino acids encoded by 1980 bp. The deletions in the Plzfb protein are mainly found within the RD2 domain and whether this has any effect on the protein function is unknown. Comparing the cDNA and amino acid sequence for the paralogues reveals that both have over 75% identity (Appendix Figure 8-1 and Figure 8-2).

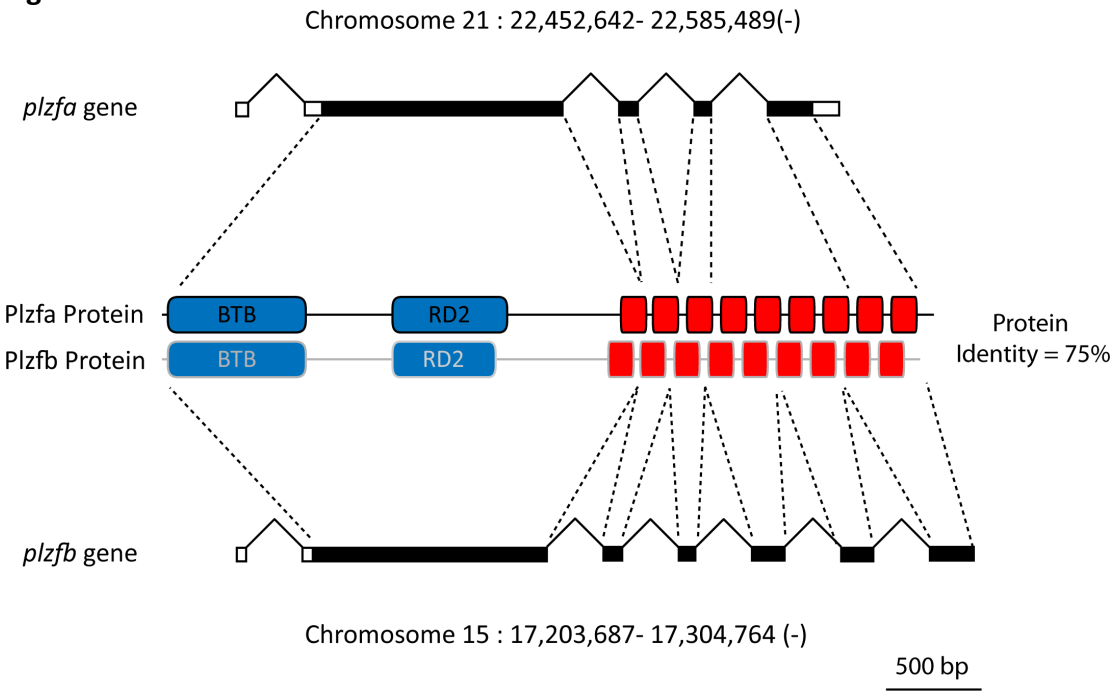
plzfa and plzfb RNA expression

I aimed to examine the expression patterns of *plzfa* and *plzfb* throughout zebrafish embryonic neurogenesis in order to determine their expression within the nervous system and observe whether the patterns overlap. To do this, I used anti-sense RNA probes against the two mRNAs for *in situ* hybridization (ISH). Because of the high degree of identity between the paralogues, I was concerned that the probes against the full length mRNA would not be specific to a particular paralogue. I therefore generated the probes from regions of least similarity between *plzfa* and *plzfb* (Appendix Figure 8-1).

Figure 4-1: Genomic location of *plzfa* and *plzfb*

Shaded boxes indicate protein-coding regions whilst the unshaded regions correspond to the untranslated regions. Exons are drawn to the scale shown, introns are not to scale.

Figure 4-1


Chromosome 21 : 22,452,642- 22,585,489(-)
plzfa gene
Plzfa Protein
BTB
RD2
Plzfb Protein
BTB
RD2
Protein
Identity = 75%
plzfb gene
Chromosome 15 : 17,203,687- 17,304,764 (-)
500 bp

In order to examine gene and protein expression from different perspectives within the developing zebrafish embryo, I used different mounting techniques for imaging. These are the Dorsal View (DV), Transverse Sections (TS) or Sagittal View (SV) as shown in the cartoons in Figure 4-2 and used throughout the thesis as indicated. A time course for *plzfa* and *plzfb* expression from the onset of neurogenesis (11 hpf) up to 44 hpf is shown in Figure 4-2a-l. The expression patterns for *plzfa* and *plzfb* closely match the patterns for the equivalent stages (11-14 hpf) previous published (Sobieszczuk et al., 2010).

At 11 hpf, *plzfa* expression is observed within the forebrain, midbrain, hindbrain and caudal neural plate at the posterior of the embryo, whilst *plzfb* is clearly expressed in the forebrain, midbrain and hindbrain (Figure 4-2a & g). As development proceeds through to 20 hpf, expression of both paralogues remains high within the developing nervous system, in particular overlapping in the anterior of the embryo (Figure 4-2b, c, h & i). Between these stages, different levels of expression were observed for the two paralogues within different presumptive rhombomeres of the developing hindbrain. This is shown in more detail in Figure 4-2m & n, where *plzfa* expression is slightly decreased in rhombomere 5 and *plzfb* has higher levels of expression within the odd-numbered rhombomeres at 20 hpf. The reason for this segmental expression was not explored any further in this project and it is possible that this expression is also dynamic between the stages imaged.

Neurogenesis is known to be most active between 24 and 44 hpf (Lyons et al., 2003), and at these stages high levels of *plzfa* transcripts could be detected within the forebrain, hindbrain and spinal cord (Figure 4-2d-f). *plzfb* expression remained high within the forebrain during this same period, but appeared to decrease elsewhere (Figure 4-2j-l). Expression of both paralogues is not restricted to the neuroepithelium and expression of *plzfa* is observed within the developing otic vesicle and pectoral fins (Figure 4-2c &f) whilst *plzfb* transcripts are observed in the developing digestive tract, such as the liver and pancreatic buds (Figure 4-2l).

Having observed expression along the anterior-posterior axis, I aimed to look in detail at the transcripts within the dorsal-ventral axis of the neural tube. I elected to compare expression of *plzfa* and *plzfb* with that of a known marker of progenitor cells, *sox3*. Based on our detailed knowledge of the hindbrain and the fact that both

paralogues are expressed there, I focussed my analysis to this region of the neural epithelium. At 24 hpf, the majority of the hindbrain is made up of progenitors, demonstrated by the widespread expression of *sox3* (Figure 4-2o), with neurons occupying the lateral regions of the neural epithelium. *plzfa* and *plzfb* expression overlap with that of *sox3*, suggesting that these genes are coexpressed within neural progenitors (Figure 4-2p & q).

At 44 hpf, the ventricular zone is in a characteristic T-shape, whilst the mantle zone contains differentiated neurons which occupy the remaining space (Lyons et al., 2003). This is reflected in the expression pattern for *sox3* (Figure 4-2r). *plzfa* and *plzfb* show a similar expression pattern to *sox3* within the ventricular zone of the hindbrain (Figure 4-2s & t). *plzfa* expression appears to be diminished in the dorsal-most region of the ventricular zone (arrows in Figure 4-2s). *plzfb* expression in the hindbrain at 44 hpf was difficult to detect in the dorsal views (Figure 4-2l), but these sections show that it clearly is expressed within the ventricular zone.

Overall, these ISHs reveal a number of interesting features about expression of the two paralogues. They show that *plzfa* and *plzfb* are both widely expressed and are highly overlapping throughout zebrafish development and that within the neural epithelium their expression is associated with neural progenitors.

Plzf protein expression

The previously characterised zfPlzf Ab (Figure 3-2) was used to analyse endogenous Plzf protein expression in the zebrafish embryo during the initial 24 hours of development (Figure 4-3a-d). Although not completely clear from these low magnification views, the antibody staining appeared to give a similar expression pattern as for the RNA for *plzfa* and *plzfb*, with high levels of expression throughout the developing nervous system. In order to determine whether or not Plzf expression occurs in differentiated neurons, I carried out IHC with both zfPlzf Ab and to detect HuC/D. Examining the hindbrain in detail shows that Plzf is expressed within the nucleus of neural progenitor cells, and is excluded from the post-mitotic, differentiated neurons (Figure 4-3e-g). As suggested by the ISH results, expression is

Figure 4-2: *plzfa* and *plzfb* expression during development

Zebrafish cartoons indicate different imaging positions used throughout this thesis. Dorsal view (DV) images are taken from embryos flat mounted on their ventral surface. Transverse sections (TS) are obtained by sectioning embryos at different points along the anterior-posterior axis. Sagittal views (SV) are taken from whole embryos mounted on their side.

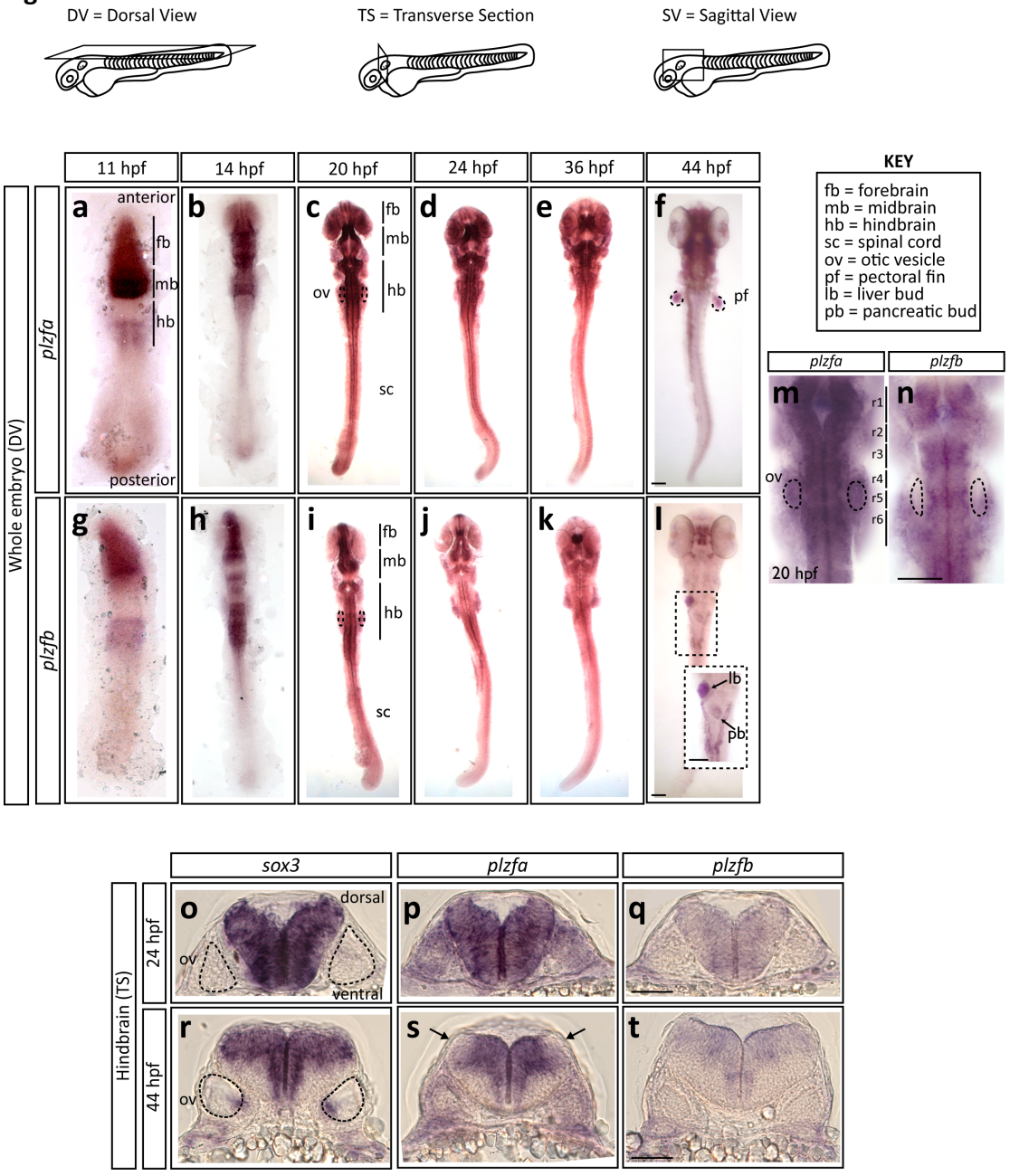
a-l: Whole mount ISH DV images taken of developing zebrafish embryos from 11 to 44 hpf for *plzfa* (a-f) and *plzfb* (g-l). Both genes show dynamic expression patterns as described within the text. Regions of the embryo are labelled as in the key. The developing otic vesicle is indicated by dotted lines (c & i), as are the emerging pectoral fins (f). The boxed region showing the digestive system is displayed at a higher magnification within the inset (l) with the liver and pancreatic bud marked by arrows.

m & n: Segmental expression of *plzfa* and *plzfb* at 20 hpf. Rhombomeres (r1-r6) are identified based on morphological characteristics such as the otic vesicle (indicated), which acts as a rough guide for r5.

a-l: Transverse sections through the hindbrain show expression of *sox3* (o & r), *plzfa* (p & s) and *plzfb* (q & t) at 24 and 44 hpf. All sections were taken where the otic vesicle could be identified (dashed line in o & r). Arrowheads indicate the dorsal-most region of the neural tube where *plzfa* expression is lacking (s).

Scale bar = 100 μm.

Figure 4-2


DV = Dorsal View
TS = Transverse Section
SV = Sagittal View
11 hpf
14 hpf
20 hpf
24 hpf
36 hpf
44 hpf
Whole embryo (DV)
plzfa
plzfb
anterior
posterior
fb
mb
hb
sc
ov
pf
lb
pb
KEY
fb = forebrain
mb = midbrain
hb = hindbrain
sc = spinal cord
ov = otic vesicle
pf = pectoral fin
lb = liver bud
pb = pancreatic bud
plzfa
plzfb
r1
r2
r3
r4
r5
r6
20 hpf
sox3
Hindbrain (TS)
24 hpf
44 hpf
dorsal
ventral

not confined to the developing neural epithelium and Plzf expression is also observed in tissues such as the otic vesicle.

Transverse sections through the hindbrain at 24 hpf reveal Plzf protein within the ventricular zone, similar to the RNA expression patterns in Figure 4-2o-q. At 44 hpf, Plzf protein expression is observed throughout the ventricular zone except for the dorsal-most region (Figure 4-3k-n). This is reminiscent of the *plzfa* RNA expression pattern shown previously (Figure 4-2s). Studying these sections at a higher magnification shows that at this stage there are some cells that express both Plzf and HuC/D (Figure 4-3k'-n'). This indicates that there are some terminally differentiated neurons that coexpress Plzf within the zebrafish hindbrain, and potential reasons for this observation will be discussed later. In all cells where Plzf was expressed it was nuclear and roughly similar levels of Plzf staining were observed across the anterior-posterior and dorsal-ventral axes within the hindbrain.

The similarity between the *plzfa* and *plzfb* RNA expression patterns (Figure 4-2) meant that it was difficult to determine whether the zfPlzf Ab was capable of detecting both proteins in the zebrafish embryo. To examine this further, I used anti-sense morpholinos (MOs) to knockdown Plzfa and Plzfb protein translation separately and analysed the resulting protein expression pattern using the antibody. Surprisingly, the *plzfa* MO was able to greatly reduce staining in zebrafish, whilst the *plzfb* MO had no clear effect on expression (Figure 4-4a-c).

In order to examine this further, I attempted to detect the Plzfa or Plzfb protein within zebrafish by performing Western Blots on MO injected embryo lysates using the zfPlzf Ab. Although I was able to clearly detect the HA-tagged Plzfa protein when transfected into HEK293 cells using this method, I did not observe any band at the expected size within lysates obtained from zebrafish embryos (Figure 4-4d). I therefore was only able to investigate the presence of Plzf protein by IHC.

There are a number of possible explanations why the *plzfa* MO abolishes the antibody staining and why the *plzfb* MO has no discernable effect. If the Plzfa protein is able to drive transcription of the *plzfb* gene then knocking down Plzfa could result in a decrease of *plzfb* transcripts and therefore Plzfb protein production. To test this, I carried out ISH for *plzfb* on embryos injected with *plzfa* MO and

Figure 4-3: Monitoring Plzf protein expression during development

a-d: Time course of zfPlzf Ab staining shows Plzf protein expression to be highly similar to that suggested by the RNA expression in Figure 4-2. The otic vesicle is marked by a dashed line (c).

e-j: Embryos were stained for Plzf and differentiated neurons, marked by HuC/D. Dorsal view at the hindbrain of 24 hpf embryos shows that Plzf is absent from differentiated neurons within the neural tube at this stage, such as at the arrowed location (e-g). Plzf can also be observed within the developing otic vesicle (e). Transverse sections of the 24 hpf (h-j) and 44 hpf (k-n) hindbrain show Plzf expression throughout the neural tube. Arrows indicate dorsal-most region where Plzf expression is absent (k-n). Examining expression at a higher magnification (box in n) reveals that some cells express both Plzf and HuC/D, indicated by arrowheads (k'-n').

Scale bar = 100 µm (a-d) and 20 µm (e-n'). Key as in Figure 4-2.

Figure 4-3

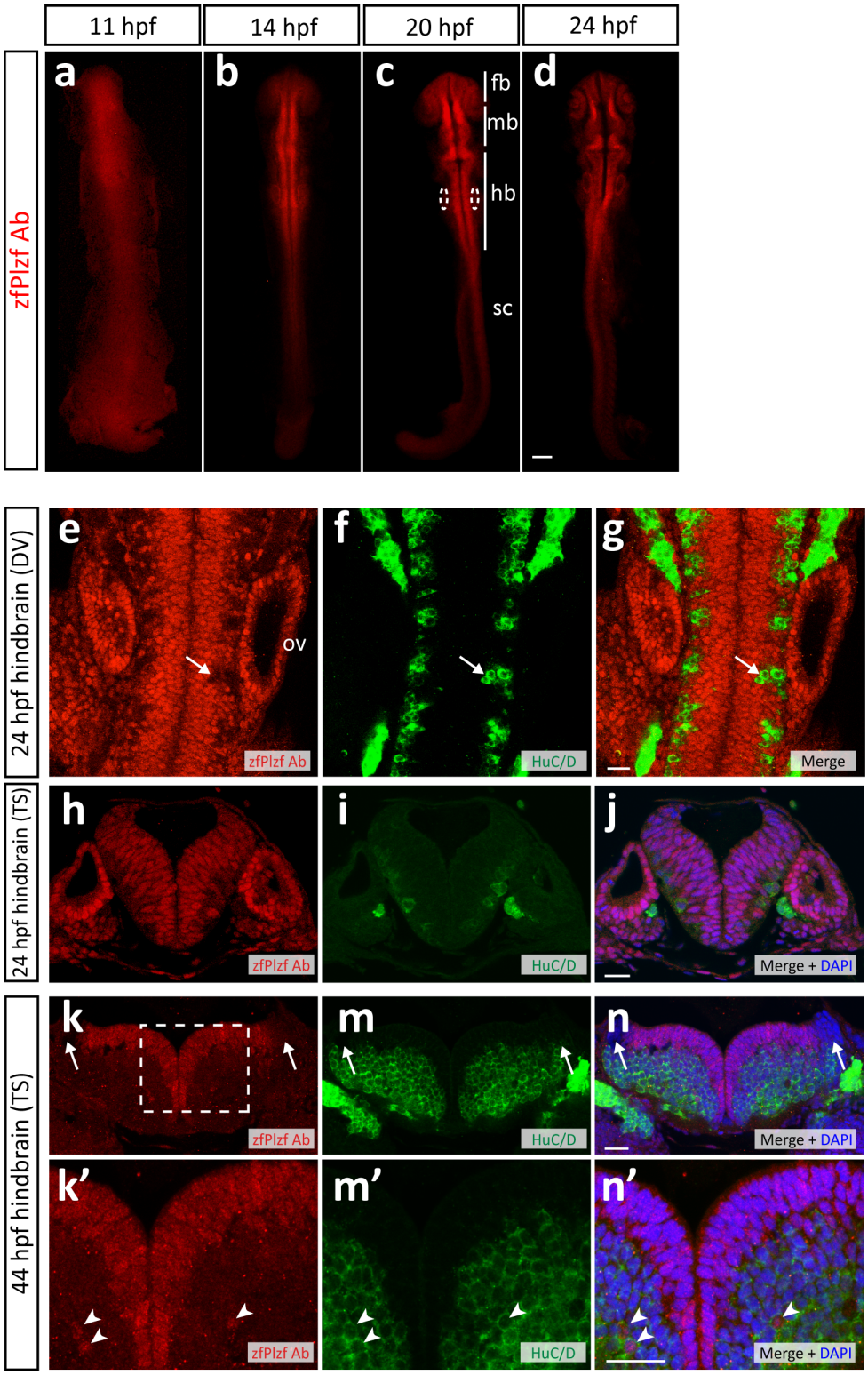

11 hpf
14 hpf
20 hpf
24 hpf
zfPlzf Ab
a
b
c
d
fb
mb
hb
sc
24 hpf hindbrain (DV)
e
f
g
ov
zfPlzf Ab
HuC/D
Merge
24 hpf hindbrain (TS)
h
i
j
zfPlzf Ab
HuC/D
Merge + DAPI
44 hpf hindbrain (TS)
k
m
n
zfPlzf Ab
HuC/D
Merge + DAPI
k'
m'
n'
zfPlzf Ab
HuC/D
Merge + DAPI

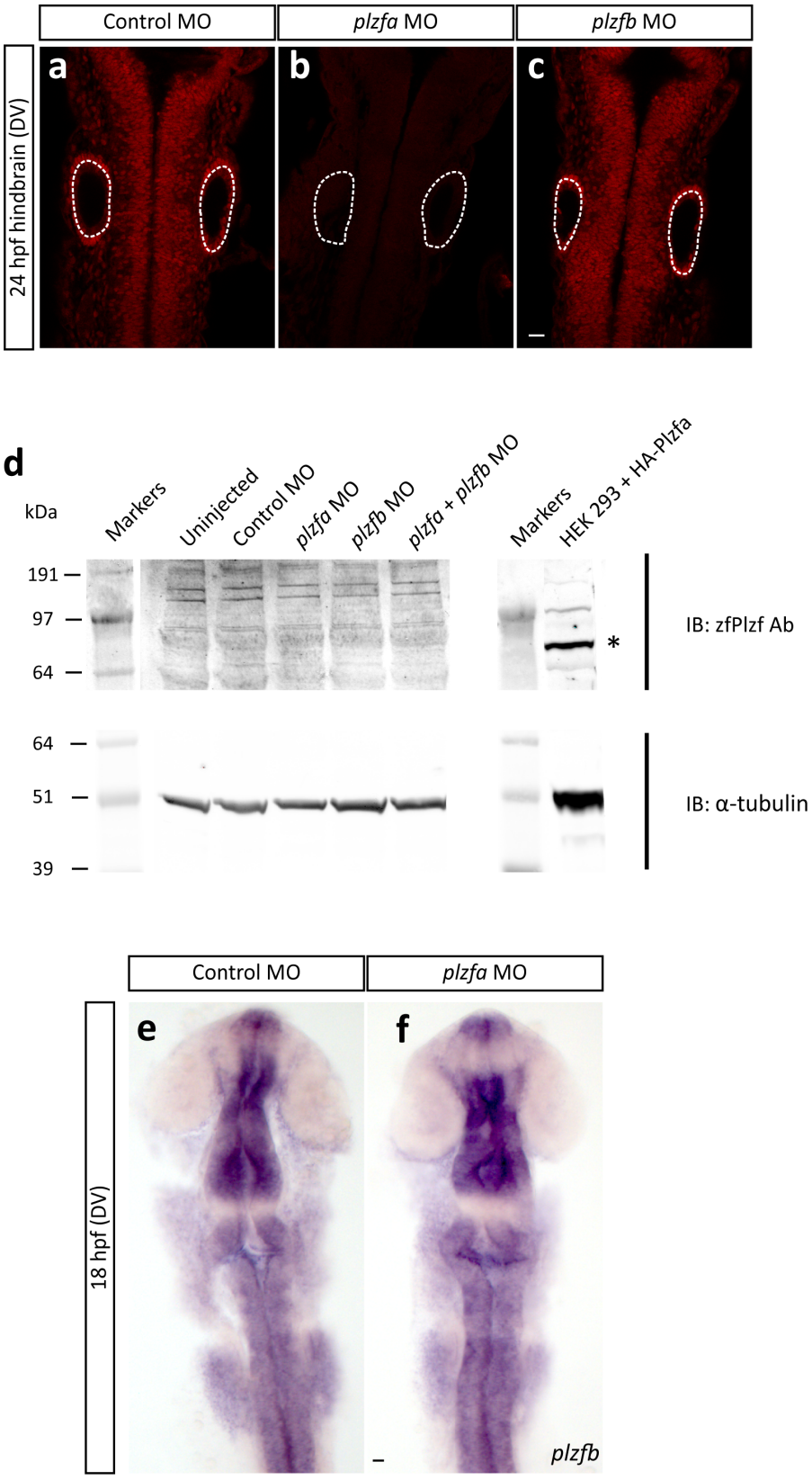
Figure 4-4: Attempts to distinguish between endogenous Plzfa and Plzfb

a-c: Monitoring Plzf protein expression in embryos injected with control (a), *plzfa* (b) or *plzfb* (c) MOs using the zfPlzf Ab. Protein expression appears entirely abolished in *plzfa* MO injected embryos and unaffected in *plzfb* MO injected at 24 hpf. Otic vesicle marked by dashed lines.

d: Attempts by western blot to detect Plzf protein. Pools of embryos that had been injected with the indicated MOs were lysed and Western Blots carried out on 30 µg protein using the zfPlzf Ab. HEK 293 cells transfected with HA-Plzfa was used as a control. A band representing HA-Plzfa can be found in the control lane and is indicated by an asterisk (*). No bands representing Plzfa or Plzfb can be detected in any of the zebrafish lysates. α-tublulin was used as a loading conrol.
Expected sizes in kDa - Plzfa = 74.9; Plzfb = 74.1; HA-Plzfa = 76.4; α-tublulin = 50.0.

e & f: *plzfb* transcript levels aren't affected by Plzfa knockdown. Representative *plzfb* ISH at 18 hpf injected with control (e) or *plzfa* (f) MOs shown.
Scale bar = 20 µm

Figure 4-4


Control MO
plzfa MO
plzfb MO
24 hpf hindbrain (DV)
a
b
c
d
kDa
Markers
Uninjected
Control MO
plzfa MO
plzfb MO
plzfa + plzfb MO
Markers
HEK 293 + HA-Plzfa
191
97
64
*
IB: zfPlzf Ab
64
51
39
IB: α-tubulin
Control MO
plzfa MO
18 hpf (DV)
e
f
plzfb

observed no significant changes in transcript levels when compared to control injected embryos (Figure 4-4e & f).

The simplest explanation is that the levels of Plzfb protein are below the threshold of detection by the zfPlzf Ab. Although not a quantitative test, the ISH results shown in Figure 4-2 suggest that *plzfb* RNA is expressed at a lower level when compared with *plzfa*. Furthermore, there is also some evidence that Plzfb protein is not efficiently translated. A well-studied regulator of protein translation is the presence of multiple open reading frames upstream of the start codon in the mRNA (Morris and Geballe, 2000). Comparing the leader sequence for both *plzfa* and *plzfb* reveals the presence of an upstream Kozak sequence in *plzfb* that would result in an out-of-frame protein being produced (Appendix Figure 8-1b). This sequence is not observed in the leader sequence for *plzfa* and could potentially result in the *plzfb* transcripts not being efficiently translated into Plzfb protein.

Taken altogether, these results show that the zfPlzf Ab can be used to detect the presence of Plzfa protein within the developing nervous system, showing that the nuclear protein is widespread throughout the ventricular zone and largely absent from differentiated cells in the hindbrain. The antibody also provides a simple method to confirm the extent of *plzfa* MO knockdown.

4.3 Neurogenesis in the zebrafish hindbrain

Having established that Plzf is present within the proliferating neural progenitor cells that reside within the ventricular zone and absent from the post-mitotic neurons in the mantle zone, I aimed to determine where in the neurogenic cascade Plzf expression is downregulated. I also aimed to investigate different markers of neurogenesis for use in Plzf functional analysis. To do this, I chose the 44 hpf zebrafish hindbrain when there is a stereotyped pattern of neuronal differentiation (Amoyel et al., 2005).

Comparison with progenitor markers

I began by using available antibodies to study progenitors and neuronal differentiation in the hindbrain. Sox2 is a known stem cell marker (Graham et al., 2003) and examining expression within the 44 hpf hindbrain reveals that, in the neural tube, it is restricted to cells in the ventricular zone (Figure 4-5b-d). At this

stage in development, the stalk of the T-shaped ventricular zone can be seen along the midline of the embryo in dorsal views (Figure 4-5a). The pan-neuronal marker, HuC/D, reveals a stereotypical pattern of differentiated neurons within the mantle zone surrounded by gaps (Figure 4-5c & f). Examining Plzf expression at the same stage reveals that as well as being expressed within the ventricular zone, Plzf protein is observed within cells that extend into these gaps within the mantle zone (Figure 4-5e-g). This analysis posed two related questions: first, why are the neuronal cell bodies arranged in this pattern; second, what are these Plzf-expressing cells outside the ventricular zone?

Previous work has shown that at this stage in development, the majority of progenitor cells are radial glial cells (Lyons et al., 2003). These mitotically active cells are in the ventricular zone and extend glial fibres to contact the pial surface at locations adjacent to the rhombomere boundaries (Trevarrow et al., 1990). I aimed to examine whether the presence of these fibres in the mantle zone could account for the pattern of HuC/D cells. I made use of a *gfap:gfp* transgenic zebrafish line that expresses GFP in the radial glial cells (Bernardos and Raymond, 2006, Kim et al., 2008).

At 28 hpf, I compared *gfap*-driven GFP expression with that of Plzf and HuC/D in the hindbrain (Figure 4-5h-k). GFP expression overlaps with that of Plzf in the neural epithelium and is absent from neuronal cells. Examining this reporter line at 44 hpf in a sagittal view shows expression within the ventricular zone and in the fibres extending from the ventricular to the pial surface, through the mantle zone (Figure 4-5l-n). The fibres can also be observed using an antibody against GFAP (Figure 4-5m).

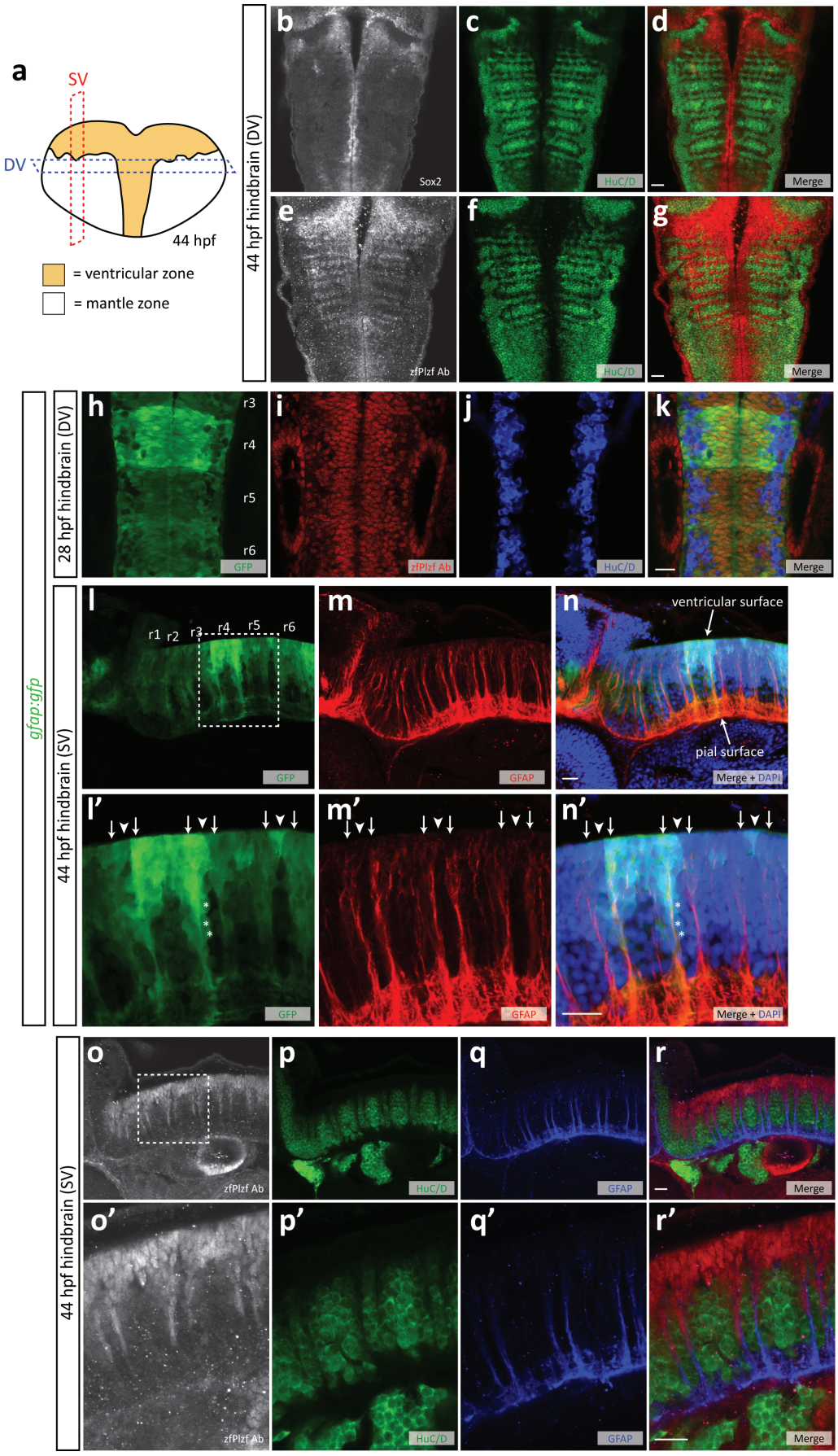
As expected, all cells of the ventricular zone express *gfap*, indicating that they are radial glial cells, and the fibres are arranged in bundles that extend from locations adjacent to rhombomere boundaries (arrows in Figure 4-5l'-n'). *gfap* expressing cells can also be observed associated with the fibres within the mantle zone (asterisks in Figure 4-5l'-n'). These cells have exited the cell cycle and are beginning to migrate along the fibres whilst undergoing the neurogenic cascade, before terminal differentiation when they reach their final position in the mantle zone. For the remainder of my thesis I shall class these cells as differentiating cells.

Figure 4-5: Plzf expression in the 44 hpf hindbrain is not restricted to cells of the ventricular zone

a: Cartoon of the 44 hpf transverse section hindbrain. Dotted lines indicate position where confocal sections were taken for DV (blue line) and SV (red line) images.
b-**g**: DV confocal images comparing Sox2 (b) expression to Plzf (e) expression within the 44 hpf hindbrain. Whilst Sox2 is restricted to the thin stripe representing the ventricular zone, Plzf expression extends into the mantle zone, present in the gaps in HuC/D expression (f).
h-n: Transgenic *gfap:gfp* embryos express GFP in neural progenitors. In the 28 hpf hindbrain, GFP expression (h) overlaps with Plzf protein (i) within the neural tube. GFP protein is not present in post-mitotic neurons (j). Rhombomeres are labelled r1 – r6 and higher levels of GFP expression are observed in r4 (h & l). 44 hpf transgenic embryos were examined along with an antibody against GFAP fibres. Most GFP expressing cells exist at the ventricular surface and extend their fibres in bundles adjacent to rhombomere boundaries and contact the pial surface (l-n). Higher magnification images of l-n are displayed in l'-n'. Some GFP positive cells can be observed within the mantle zone attached to the fibres (examples marked by asterisk in l' & n'). Arrows indicate regions where the fibres are extending and arrowheads mark rhombomere boundaries (l'-n').
n-r: Plzf is expressed in cells within the mantle zone. A sagittal view of the 44 hpf hindbrain shows streams of cells expressing Plzf (o) migrating along the glial fibres (q). These streams represent the same cells as in the stripes in the dorsal view (e). Higher magnification images of o-n are displayed in o'-n'.

Embryo orientations as described in Figure 4-2 cartoons.
Scale bar = 20 µm

Figure 4-5


a
SV
DV
44 hpf
= ventricular zone
= mantle zone
44 hpf hindbrain (DV)
b
Sox2
c
HuC/D
d
Merge
e
zfPlzf Ab
f
HuC/D
g
Merge
28 hpf hindbrain (DV)
h
r3
r4
r5
r6
GFP
i
zfPlzf Ab
j
HuC/D
k
Merge
gfap:gfp
44 hpf hindbrain (SV)
l
r1 r2 r3 r4 r5 r6
GFP
m
GFAP
n
ventricular surface
pial surface
Merge + DAPI
l'
GFP
m'
GFAP
n'
Merge + DAPI
44 hpf hindbrain (SV)
o
zfPlzf Ab
p
HuC/D
q
GFAP
r
Merge
o'
zfPlzf Ab
p'
HuC/D
q'
GFAP
r'
Merge

Examining Plzf expression from the sagittal view shows that Plzf is expressed in at least some of these differentiating cells (Figure 4-5o). From these images it appears that it is not only the glial fibres that are preventing the neuronal cell bodies from occupying space in the mantle zone, but it is the presence of the migrating cells themselves (Figure 4-5o-r). This analysis reveals that Plzf expression is not restricted to the ventricular zone, but also occurs in differentiating cells within the mantle zone and that the presence of these cells contributes to the stereotypical pattern of neuronal cell bodies observed within the hindbrain.

Comparison with neurogenic cascade markers

Evidence that these migrating cells are actively undergoing neurogenesis come from the knowledge that various proneural genes are expressed adjacent to the rhombomere boundaries (Amoyel et al., 2005, Gonzalez-Quevedo et al., 2010). Proliferating progenitor cells contain low, or oscillating, levels of the proneural gene *neurog1*, which remains at high levels in cells that have initiated neurogenesis (Kim et al., 1997). *neurod4* is expressed in differentiating, migrating cells prior to terminal differentiation (Roztocil et al., 1997, Wang et al., 2003). At 44 hpf, expression of both genes was observed at locations flanking the rhombomere boundaries, outside of the Sox2 expressing cells (Figure 4-6a-c). Some cells expressing both *neurog1* and *neurod4* are observed near the ventricular surface and cells expressing only *neurod4* are located in more basal positions.

I made attempts to combine the ISH for these genes with IHC using zfPlzf Ab, but couldn't detect a signal for Plzf protein. Instead, I carried out a double ISH for *plzfa* and *neurod4* (Figure 4-6d-f). Similar to the Plzfa protein, *plzfa* transcripts can be detected in cells both within the ventricular zone and also in the streams of cells on the glial fibres (Figure 4-6d). Some coexpression is observed with *neurod4* (arrows in Figure 4-6f'), but the majority of expression is observed apical to the *neurod4* expressing cells. A limitation of this double ISH technique is that development of one probe can block the development of the next and therefore it is difficult to reliably detect transcripts with overlapping expression (Lauter et al., 2011). For this reason I was unable to clearly show coexpression of *plzfa* and *neurog1*, but the results suggest this to be the case.

As far as I am aware, there are no antibodies that work in zebrafish that recognise markers specifically expressed in these differentiating cells. Observations by Jordi Cayuso (unpublished data) noted that the *prox1* gene was expressed in the zebrafish neural epithelium. Studies in other vertebrates have determined that Prox1 is expressed downsteam of proneural genes and defines a transitory cell state between progenitor and differentiated neuron (Torii et al., 1999, Misra et al., 2008). Examining expression of this protein by IHC in the zebrafish hindbrain shows expression within cells of the neurogenic stripes outside of the ventricular zone and absent from differentiated neurons (Figure 4-6g-j). Comparison with Plzf protein (Figure 4-6g) shows that expression of the two proteins largely overlap in these stripes. A sagittal view shows the Prox1 positive cells associated with the glial fibres (Figure 4-6k & l).

In summary, I have shown that Plzf expression is not restricted to progenitors of the ventricular zone but is also expressed within cells actively undergoing neurogenesis in the 44 hpf hindbrain. These differentiating cells initially express *neurog1* as they leave the ventricular zone and later express *neurod4,* by which point *plzfa* expression is severely decreased. During this time they are also expressing Prox1. Radial glial progenitors project their glial fibres from locations adjacent to rhombomere boundaries and it is along these fibres that the differentiating cells are present. The presence of these cells prevents neuronal cell bodies from occupying this space, therefore leading to a highly stereotypical pattern of differentiated neurons in the mantle zone. A summary model of the expression data is shown in Figure 4-7.

4.4 Analysis of the function of Plzfa and Plzfb

In order to investigate the function of Plzfa and Plzfb within zebrafish, I used the previously described translation-blocking MOs to knockdown expression. The MOs are injected into a *tp53* null background to prevent the known off-target effects related to apoptosis (Robu et al., 2007, Gerety and Wilkinson, 2011). Injection of the *plzfa* MO resulted in a delay in development of the embryo of around 4 hours, clearly identifiable at stages before 24 hpf by counting the number of somites. I confirmed that the 4 hour delay is consistent for later stage embryos by

Figure 4-6: Plzf expression overlaps with neurogenic markers

a-c: Double fluorescent ISH followed by IHC was carried out at 44 hpf for *neurog1* (a), *neurod4* (b) and GFAP (c). The neurogenic genes are expressed at locations defined by glial fibres. Dashed line indicates dorsal surface of the embryo.

d-f: Double fluorescent ISH showing *plzfa* (d) and *neurod4* (e) with the glial fibres (f). *plzfa* is expressed within the ventricular zone and in cells undergoing neurogenesis.

g-j: Within the 44 hpf hindbrain, Prox1 protein (h) is expressed within the stripes of differentiating cells, is absent from differentiated neurons (j) and this expression partially overlaps with Plzf (g & i).

k-l: Confirmation that Prox1 (k) is expressed within differentiating cells present on the glial fibres (l) observed from a sagittal view.

Embryo orientations as described in Figure 4-2 cartoons.

Scale bar = 20 μm

Figure 4-6

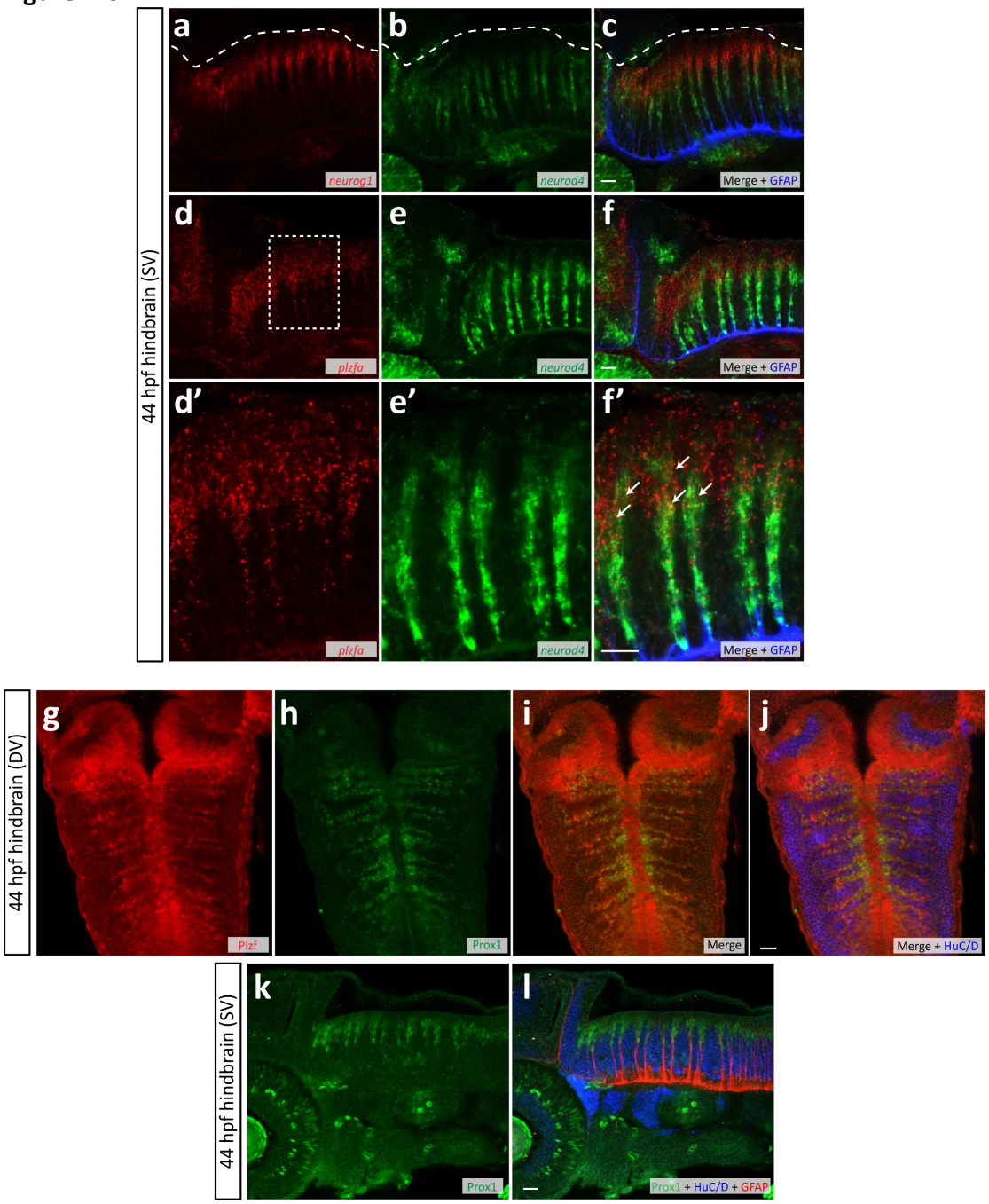
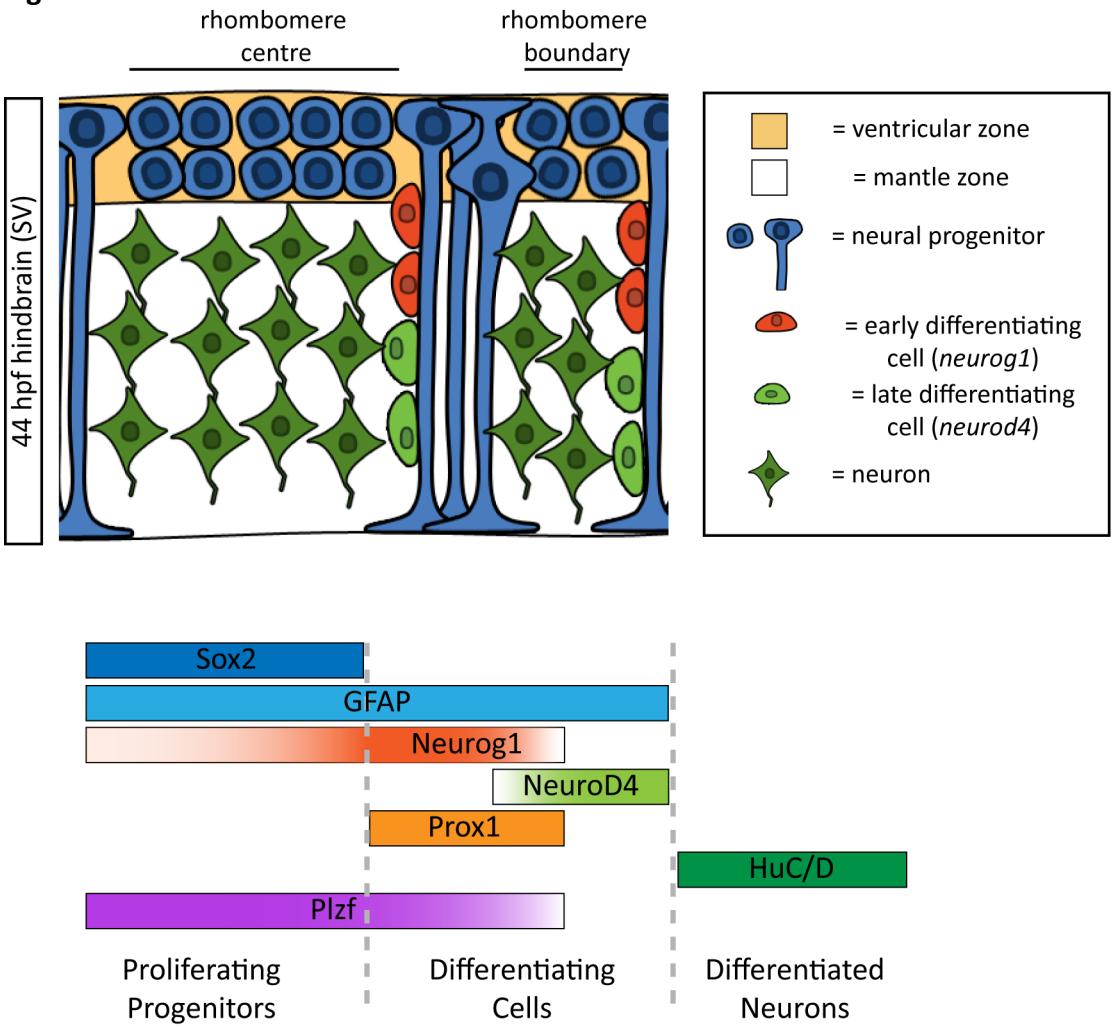

44 hpf hindbrain (SV)
a
neurog1
b
neurod4
c
Merge + GFAP
d
plzfa
e
neurod4
f
Merge + GFAP
d'
plzfa
e'
neurod4
f'
Merge + GFAP
44 hpf hindbrain (DV)
g
Plzf
h
Prox1
i
Merge
j
Merge + HuC/D
44 hpf hindbrain (SV)
k
Prox1
l
Prox1 + HuC/D + GFAP

Figure 4-7: Model of Plzf expression within the hindbrain

Summary model showing expression of genes within the 44 hpf zebrafish hindbrain described in the results. Plzf is expressed both within proliferating progenitors of the ventricular zone and within some differentiating cells but is absent from differentiated neurons within the mantle zone. The presence of the differentiating cells at precise locations contributes to the pattern of neurons in the hindbrain. Locations of the rhombomere boundary and centres are marked in the model.

Figure 4-7


rhombomere centre
rhombomere boundary
44 hpf hindbrain (SV)
= ventricular zone
= mantle zone
= neural progenitor
= early differentiating cell (neurog1)
= late differentiating cell (neurod4)
= neuron
Sox2
GFAP
Neurog1
NeuroD4
Prox1
HuC/D
Plzf
Proliferating Progenitors
Differentiating Cells
Differentiated Neurons

making use of a transgenic line with which I staged the embryos based upon the migration of the lateral line primordium (Appendix Figure 8-3). In order to accurately compare morphant embryos, I took batches at the same point in time (time-matched) and also at the same stage (stage-matched) in development by allowing the delayed embryos to develop for longer.

Phenotype relating to a loss of progenitors in the hindbrain

I observe a consistent phenotype upon knockdown of both Plzfa and Plzfb together. I compared expression of HuC/D in the 44 hpf hindbrain of control embryos (Figure 4-8a & e) against stage-matched embryos injected with *plzfa* MO (Figure 4-8b & f), *plzfb* MO (Figure 4-8c & g) and those injected with both *plzfa* and *plzfb* MO (Figure 4-8d & h). The normally restricted pattern of differentiated neurons at this stage appears slightly disrupted in the *plzfa* morphants, unchanged in the *plzfb* morphants and significantly disrupted in the double morphants. This disruption is observed both from a dorsal view (Figure 4-8a-d) and a sagittal view (Figure 4-8e-h) and it appears that, in the affected embryos, neurons invade the the gaps. To confirm that the phenotype is not caused by the developmental delay, I checked embryos that were time-matched and observed the same phenotype (Appendix Figure 8-4).

A potential explanation for this result is that Plzfa and Plzfb knockdown results in depletion of differentiating cells allowing expansion of the domains of neurons within the mantle zone. Two genes expressed within the differentiating cells at this stage are *sox3* and *neurod4* (Figure 4-9a-d). For both cases, I observed a loss of expression within the neurogenic stripes, consistent with a loss of differentiating cells within the hindbrain. As well as within the cells undergoing neurogenesis, *sox3* is strongly expressed within the ventricular zone. This expression did not appear affected in the double morphant embryos.

A loss of differentiating cells could be a consequence of a depletion of progenitor cells in the ventricular zone. I therefore compared expression of Sox2 protein between control and double morphant embryos (Figure 4-9e & f). I did not observe a consistent difference in either the number of cells expressing Sox2 or the intensity of expression. An indirect method of detecting progenitor cells is to

Figure 4-8: Knocking down Plzfa and Plzfb together disrupts the pattern of differentiated neurons in the hindbrain

a-d: Confocal slices of representative control MO (a), *plzfa* MO (b), *plzfb* MO (c) and double morphant embryos (d) showing HuC/D cells in the 44 hpf hindbrain. Higher magnification images are shown (a'-d'). Arrows (a') indicate the gaps in HuC/D expression, which are slightly disrupted in *plzfa* MO (b', 10/10) and significantly disrupted in double morphant embryos (d', 17/18). No phenotype was observed in control (a', 0/8) or *plzfb* (c', 0/5) injected embryos.

e-h: Sagittal views of injected embryos. Gaps between neurons (arrows in e) are disrupted in *plzfa* MO (f) and double morphant (h) embryos.

An asterisk (*) indicates that the embryos were developed for a further 4 hours.

Scale bar = 20 μm

Figure 4-8

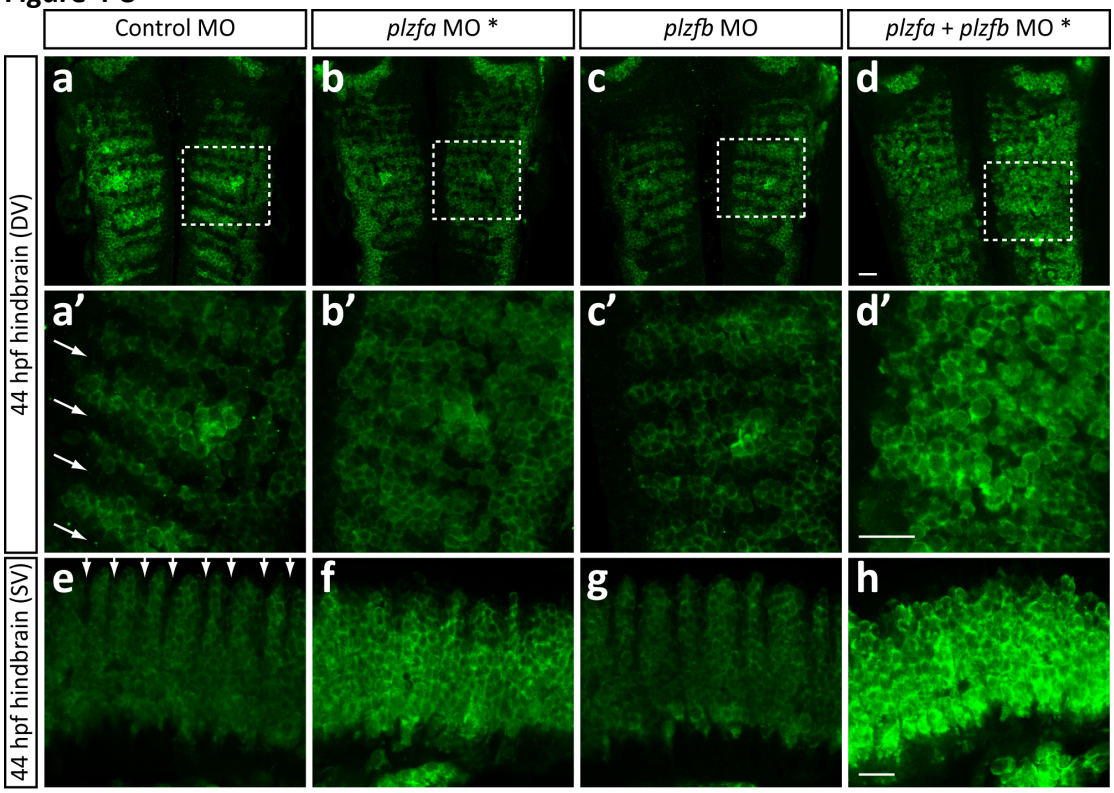
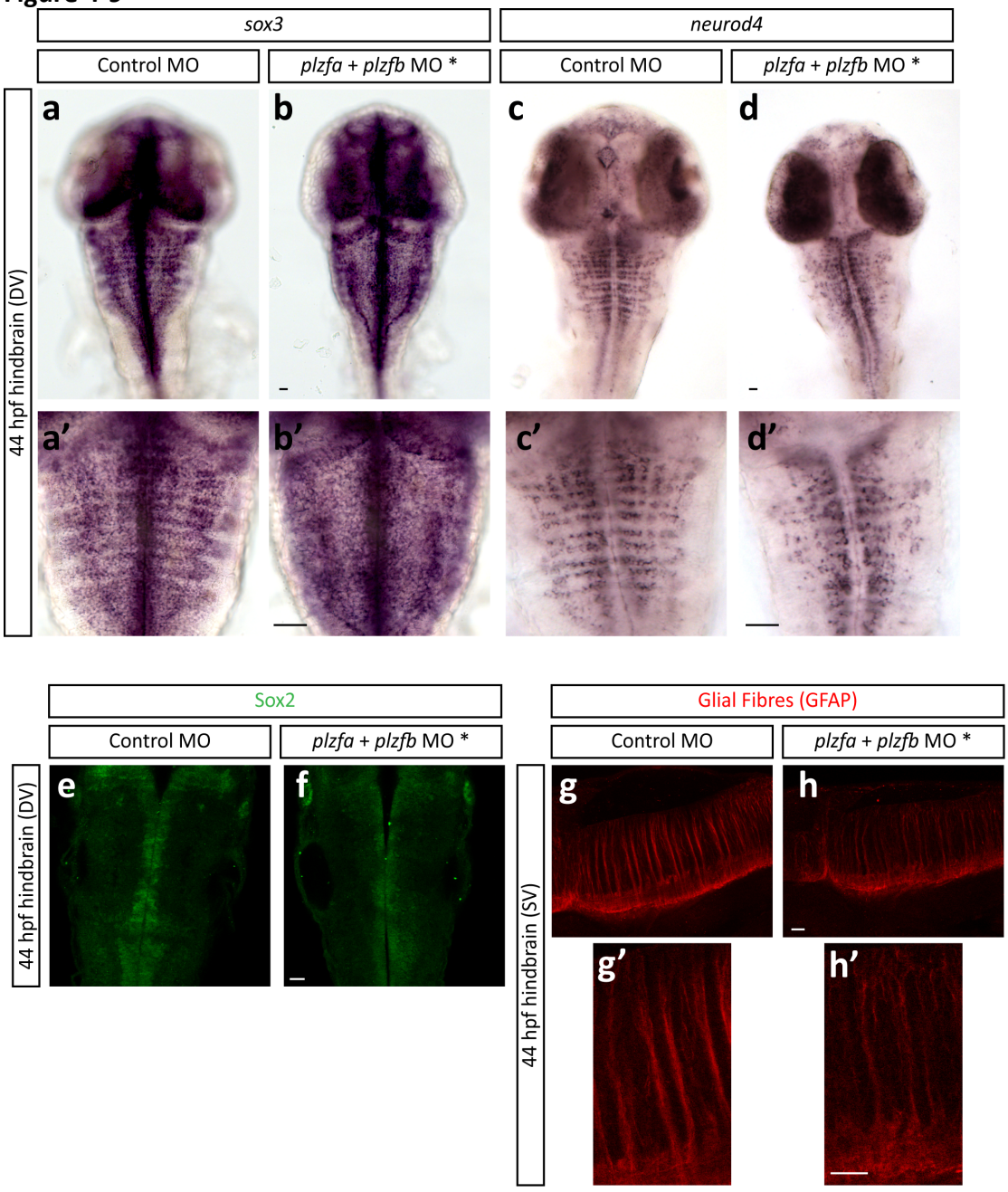

Control MO
plzfa MO *
plzfb MO
plzfa + *plzfb* MO *
44 hpf hindbrain (DV)
44 hpf hindbrain (SV)
a
b
c
d
a'
b'
c'
d'
e
f
g
h

Figure 4-9: Loss of progenitors in the 44 hpf double morphant

a-d: Loss of differentiating cells in the double morphants. Both *sox3* (a) and *neurod4* (c) are expressed within the cells undergoing neurogenesis at 44 hpf, seen as stripes within the dorsal view shown. In the double morphants (b & d), expression of both genes appears to be lost within these stripes (5/5 & 6/8 for *sox3* and *neurod4* respectively). Higher magnification images are shown (a'-d').

e-h: Examining progenitor markers reveals mixed results. No consistent changes in Sox2 protein expression were observed in the double morphant (f, n=15) compared to control embryos (e, n=15). Examining the glial fibres reveals that thinner bundles of fibres are running through the mantle zone in the double morphants (h, 11/12). An asterisk (*) indicates that the embryos were developed for a further 4 hours. Scale bar = 20 μm

Figure 4-9


sox3
neurod4
Control MO
plzfa + plzfb MO *
Control MO
plzfa + plzfb MO *
44 hpf hindbrain (DV)
a
b
c
d
a'
b'
c'
d'
Sox2
Glial Fibres (GFAP)
Control MO
plzfa + plzfb MO *
Control MO
plzfa + plzfb MO *
44 hpf hindbrain (DV)
e
f
44 hpf hindbrain (SV)
g
h
g'
h'

analyse expression of the glial fibres projected by these cells. I find that there are thinner bundles of fibres running through the mantle zone in the double morphants (Figure 4-9g & h).

The results obtained so far are suggestive of a depletion of neural progenitors upon loss of function of both Plzfa and Plzfb. One possible cause for this is premature differentiation. I therefore looked at whether there were any changes in neuronal differentiation at an earlier time point in development (30 hpf). I used the EphA4 antibody in order to distinguish between the different rhombomeres and counted the number of HuC/D expressing neurons within rhombomeres 3 – 5 (Figure 4-10a-c). Quantification revealed no significant differences in the number of neurons between the different morphants. I also analysed expression of *neurod4* at this stage and was unable to observe any differences between the control and double morphant embryos (Figure 4-10d-e).

An alternative explanation for a loss of progenitors could be a decrease in cell proliferation in the morphants. I made attempts to observe changes in proliferation in the morphants at 30 and 44 hpf, but the MO-induced delay in development made it difficult to obtain conclusive results (*data not shown*). Instead, I carried out overexpression experiments to see whether ectopic Plzfa would alter cell proliferation. I injected Plzfa RNA at the 2-cell stage to overexpress it in half of the embryo, together with H2B-Citrine as a reporter (Figure 4-10f-k). A more detailed description of the construct used for overexpression can be found in Chapter 3.4. I confirmed that overexpression induces a phenotype by staining for Islet1 protein, observing a decrease in the number of expressing cells in the injected half of the embryo (Sobieszczuk et al., 2010). Proliferation was assessed by immunostaining with phospho-histone H3 (pH3), which marks cells undergoing mitosis. I observed no significant changes in the number of proliferating cells between the injected and uninjected halves of the embryos.

4.5 Revisiting the model with Btbd6a

Previous work within the lab outlined a model for primary neurogenesis where the adapter protein Btbd6a removes Plzfa from the nucleus and targets it for

degradation (Figure 1-8) (Sobieszczuk et al., 2010). The previous evidence that Btbd6a causes the degradation of Plzfa in zebrafish came from experiments where both proteins were overexpressed. Using the zfPlzf Ab, I aimed to confirm that Btbd6a is able to target the endogenous Plzfa protein. Injecting Myc-tagged Btbd6a RNA into the embryo resulted in lower levels of endogenous Plzf protein in cells that contained exogenous Btbd6a (Figure 4-11).

Figure 4-10: Knockdown has no discernable effect on either neuronal differentiation or cell proliferation

a-c: No changes in the number of differentiated neurons in the double morphants. Confocal images showing control MO (a) or double MO (b) injected embryos at 30 hpf immunostained for EphA4 (red) and HuC/D (green). Positions of rhombomeres 3 – 5 (r3-r5), revealed by EphA4 staining, are indicated. Number of HuC/D expressing cells within r3 – r5 were counted and no significant differences between the different morphants was noted (c). Error bars represent standard error of the mean; statistical significance was determined by one-way ANOVA test. Control: n = 4, *plzfa* MO: n = 6, *plzfb* MO: n = 3, *plzfa* + *plzfb* MO n = 8.

d & e: No differences in *neurod4* expression was observed. ISH showing *neurod4* expression comparing embryos injected with control (d, n=6) or *plzfa* + *plzfb* (e, n=8) MO. Higher magnification images are shown (d' & e').

f-k: No changes in cell proliferation was observed with Plzfa was overexpressed. 1 cell at the 2 cell-stage was injected with RNA encoding H2B-Citrine and Myc-Plzfa (f). Embryos were immunostained for: Citrine (g), to show the side of the embryo injected; Phospho-Histone H3 (pH3), to show the number of actively proliferating cells (h); and Islet1 (j) to confirm that the overexpressed construct can induce a phenotype. Maximum projections from a representative embryo are shown (g-j). The number of cells expressing pH3 was counted (k) and no consistent changes were observed when comparing the injected side to the uninjected side. This is shown in a graphic format (k).

An asterisk (*) indicates that the embryos were developed for a further 4 hours.

Scale bar = 20 μm

Figure 4-10

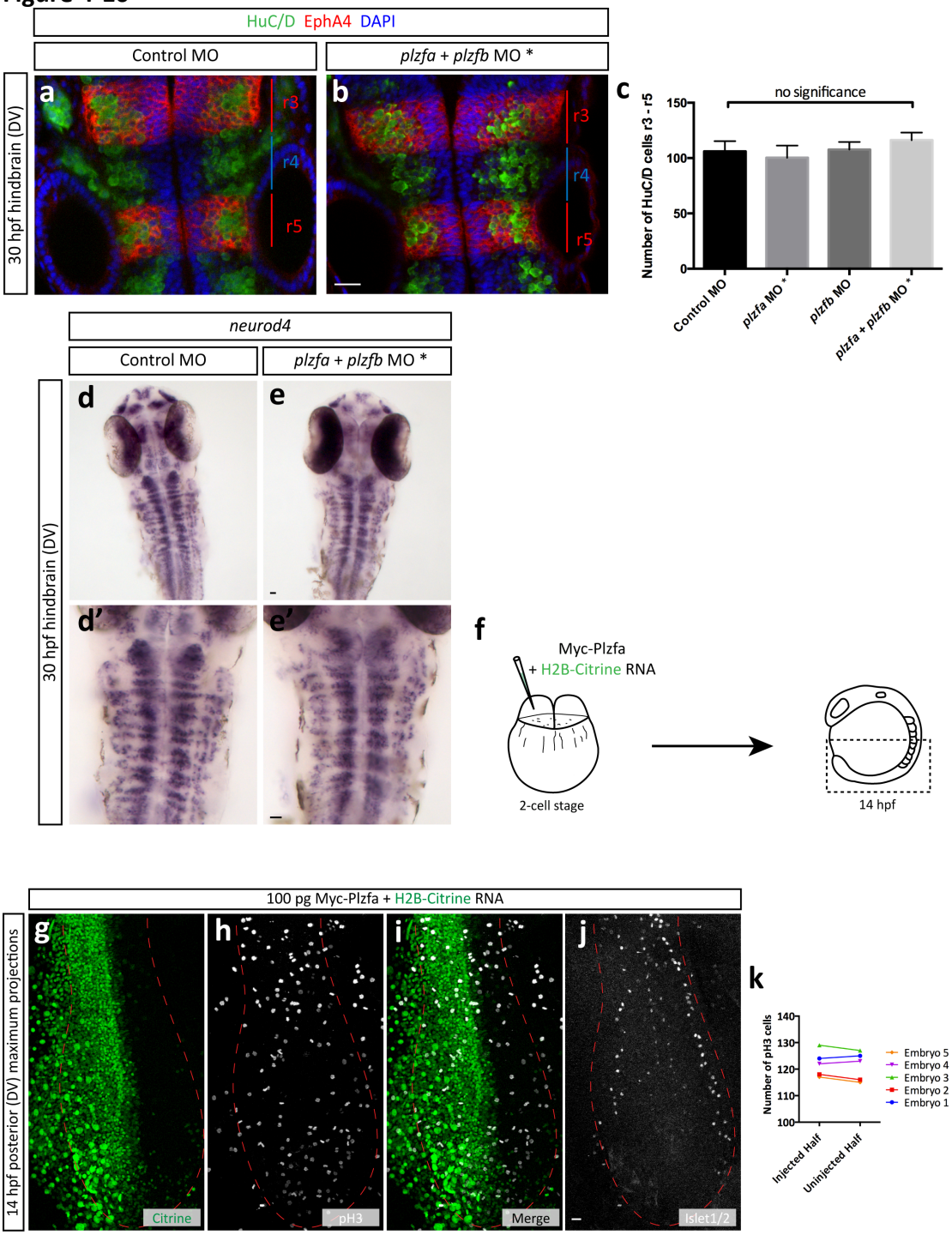
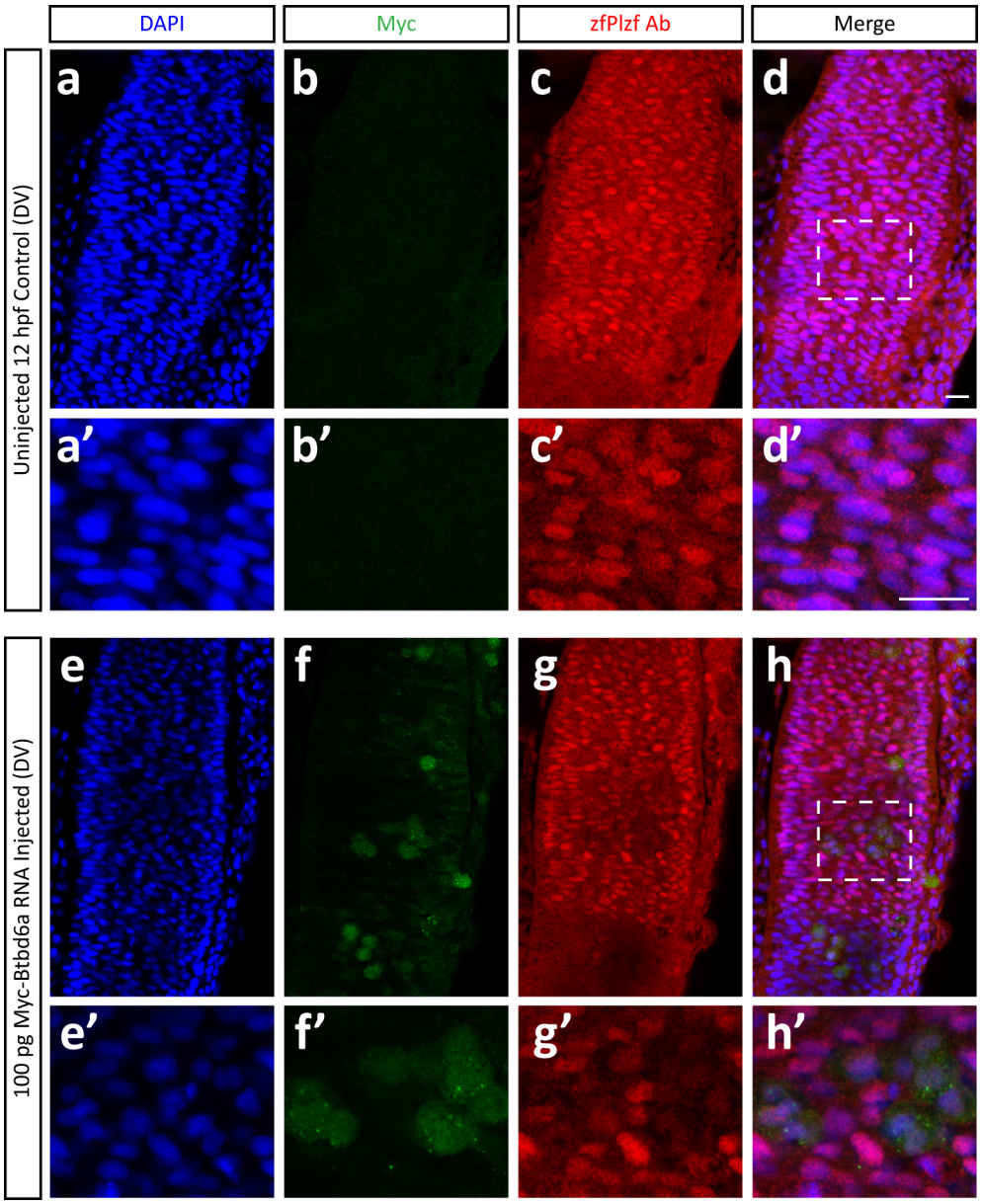

HuC/D EphA4 DAPI
Control MO
plzfa + plzfb MO *
30 hpf hindbrain (DV)
a
b
r3
r4
r5
c
no significance
Number of HuC/D cells r3 - r5
150
100
50
0
Control MO
plzfa MO *
plzfb MO
plzfa + plzfb MO *
neurod4
Control MO
plzfa + plzfb MO *
30 hpf hindbrain (DV)
d
e
d'
e'
f
Myc-Plzfa
+ H2B-Citrine RNA
2-cell stage
14 hpf
100 pg Myc-Plzfa + H2B-Citrine RNA
14 hpf posterior (DV) maximum projections
g
h
i
j
Citrine
pH3
Merge
Islet1/2
k
Number of pH3 cells
140
130
120
110
100
Injected Half
Uninjected Half
Embryo 5
Embryo 4
Embryo 3
Embryo 2
Embryo 1

Figure 4-11: Injecting Btbd6a RNA causes degradation of endogenous Plzf

Dorsal view of an uninjected, wildtype 12 hpf embryo showing endogenous, nuclear Plzf expression (a-d). Comparison with the same region of the embryo that has been injected with Myc-tagged Btbd6a RNA (e-h). Plzf appears to be largely absent from cells that contain the Myc tag.

Scale bar = 20 µm

Figure 4-11


DAPI
Myc
zfPlzf Ab
Merge
Uninjected 12 hpf Control (DV)
100 pg Myc-Btbd6a RNA Injected (DV)
a
b
c
d
a'
b'
c'
d'
e
f
g
h
e'
f'
g'
h'

4.6 Discussion

plzfa and plzfb are expressed in the ventricular zone

In this work I have carried out a detailed analysis of expression of Plzf within the zebrafish hindbrain. Building on earlier work that found *plzfa* transcripts throughout regions of the embryo associated with primary neurogenesis (Sobieszczuk et al., 2010), I show that both paralogues are coexpressed within neural progenitors of the hindbrain during the first 44 hours of development. Although not examined in detail, at low magnification the ISH patterns suggest that both genes are also expressed throughout other regions of the developing CNS.

I used the zfPlzf Ab raised against the Plzfa BTB domain in order to monitor protein expression. As described above, this antibody could reliably detect endogenous Plzfa protein but not Plzfb during development of the zebrafish embryos. Plzf protein was observed within the nucleus of the neural progenitor cells throughout the first 2 days of embryogenesis. Previously published work has shown that Plzf protein is expressed throughout the neural plate during murine, chicken and human development, suggesting that Plzf expression is conserved across species (Cook et al., 1995, Tailor et al., 2013, Gaber et al., 2013).

Detailed analysis of Plzf protein expression revealed that it was also found in the nucleus of a small number of terminally differentiated, HuC/D positive neurons within the mantle zone at 44 hpf. The previous model of Plzf function within zebrafish shows that Plzf protein is being removed from cells as they differentiate (Sobieszczuk et al., 2010). It is possible this result is indicative of newly differentiated neurons, that haven't yet removed all of Plzf from the nucleus. Analysis of Plzf expression within the chicken spinal cord found that the protein was present within a subset of differentiated interneurons (Gaber et al., 2013). It is therefore possible that a similar phenomenon is occurring in the zebrafish hindbrain and Plzf expression is maintained within certain neurons. Further analysis that compares Plzf expression with markers specific to a class of neurons would need to be carried out to understand more about this expression.

Published work showed that Btbd6a transported Plzf from the nucleus to the cytoplasm and caused its subsequent degradation by forming a complex with Cul3

(Sobieszczuk et al., 2010). Without the tools to detect endogenous Plzfa protein, this work was based on mosaic overexpression of epitope tagged versions of both Plzfa and Btbd6a in the embryo and subsequently determining the localisation of the two proteins. Access to the zfPlzf Ab meant that I was able to detect endogenous Plzf protein and therefore I confirmed ectopic Btbd6a is able to target Plzf for degradation. In my analysis of the endogenous protein, I never observed it within the cytoplasm, likely because the levels of Plzfa are below the threshold of detection by the zfPlzf Ab.

Pattern of neurogenesis in the 44 hpf hindbrain

Previous work in zebrafish had shown that by 44 hpf the majority of progenitor cells within the hindbrain have a radial glial morphology (Lyons et al., 2003). These cells extend their fibres at locations flanking the rhombomere boundaries, forming a structure known as the glial curtain (Trevarrow et al., 1990). Furthermore, these same locations are associated with cells undergoing neurogenesis, revealed by expression of proneural genes *neurog1* and *neurod4* (Amoyel et al., 2005, Gonzalez-Quevedo et al., 2010). These differentiating cells are known from studies in other systems to be migrating towards their position of terminal differentiation whilst undergoing the neurogenic cascade (Roztocil et al., 1997). Glial fibres provide scaffolding for the migration of differentiating neurons (Rakic, 2003).

Taking this information together, I show that hindbrain progenitors exiting the ventricular zone express high levels of *neurog1* at apical locations on the glial fibres. They presumably migrate along the fibres, reaching more basal locations when they express *neurod4* and downregulate *neurog1*. These two proneural genes are no longer expressed as terminal differentiation occurs, marked by the presence of HuC/D protein. The position of the differentiating cells along the fibres in the mantle zone prevent the neuronal cell bodies from occupying these locations and hence lead to the stereotypical pattern of neurons in the hindbrain mantle zone at 44 hpf.

It is unclear from my analysis why radial glial fibres are only observed flanking rhombomere boundaries. The two possibilities are that the radial glial cells are

everywhere but only extending fibres at these locations, or the progenitor cells at the rhombomere centres and boundaries display a different morphology. The observation that *gfap* driven GFP expression is in all cells of the ventricular zone suggests they all contain some glial character.

There is an association between cells undergoing neurogenesis and the location of the glial fibres in the zebrafish hindbrain (see Figure 1-5). Previous work has shown that FGF signalling from a class of neurons within the mantle zone prevents neurogenesis from occurring within segment centres and that these cells initiate neurogenesis if FGF signalling is disrupted (Gonzalez-Quevedo et al., 2010). Unpublished work has revealed that disrupting FGF signalling also results in ectopic glial fibres within the rhombomere centres (Rosa Gonzalez-Quevedo, unpublished results). Potential explanations for this result are that the progenitor cells at the segment centres are becoming radial glial cells and extending their own fibres, or alternatively the ectopic fibres could be originating from the neurogenic progenitors flanking the boundaries. Further analysis using these tools could expand our understanding of the different neural progenitors present in the hindbrain.

Plzf is expressed within differentiating cells

Analysis of Plzf protein expression at 44 hpf showed that as well as being present within neural progenitor cells, Plzf is expressed in differentiating cells at the early stages of neuronal differentiation. A lack of antibodies meant that I was initially unable to compare Plzf protein expression with markers specific for different stages of neurogenesis. I therefore compared *plzfa* RNA by carrying out double ISH analysis with proneural genes. Although difficult to precisely position the point at which *plzfa* expression is switched off during neurogenesis, the evidence suggests that it is coexpressed with *neurog1* and downregulated once *neurod4* is expressed.

I carried out some preliminary characterisation of Prox1 protein expression within the zebrafish hindbrain. Similar to what has been reported for other systems (Torii et al., 1999, Misra et al., 2008), Prox1 was found to be expressed in cells undergoing neurogenesis but absent from both progenitors and terminally differentiated neurons. I observed coexpression of Plzf and Prox1 within the

hindbrain, consistent with the idea that both the Plzfa transcript and Plzf protein are present within differentiating cells.

Evidence that Plzfa and Plzfb function together to maintain progenitors

The protein similarity between Plzfa and plzfb and the observation that both paralogues are expressed in progenitors of the developing hindbrain is suggestive that are they are functionally redundant. I therefore aimed to remove the function of both proteins simultaneously using MOs. When I do this, I observe a striking altered pattern of differentiated neurons within the 44 hpf hindbrain, whilst the single knockdown lead to no change or a mild effect.

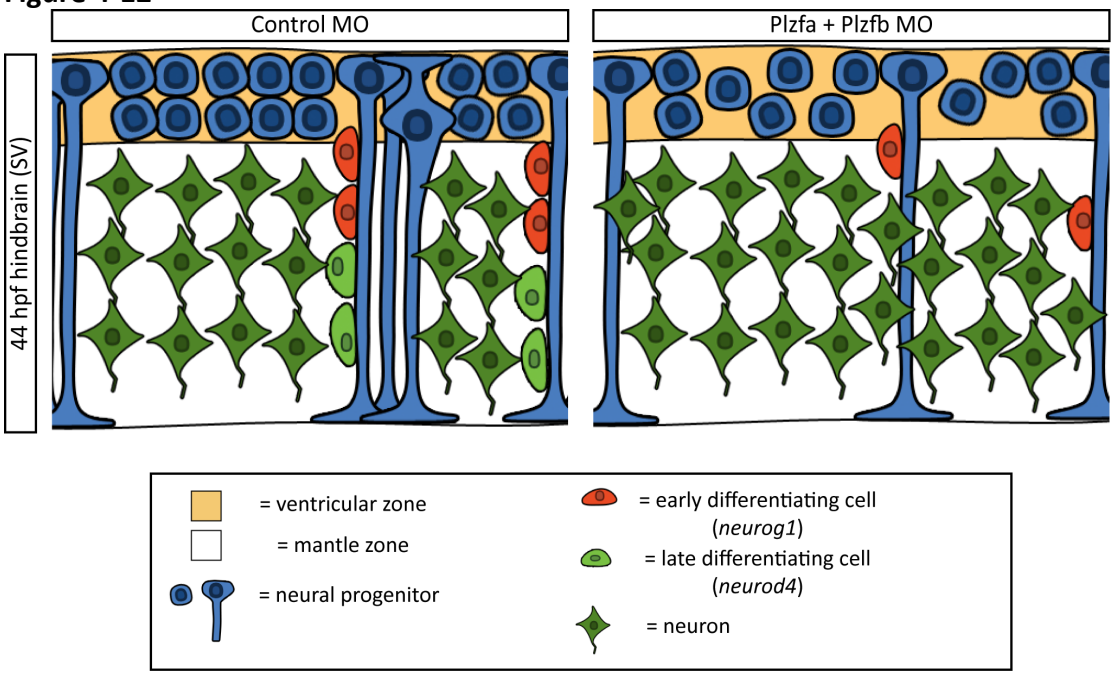
This phenotype was attributed to a loss of differentiating cells within the mantle zone, which, as previously discussed, are involved in the stereotypical patterning of differentiated neurons in the hindbrain. This was confirmed by observing a decrease of cells expressing neurogenic markers at 44 hpf. The likely explanation for a loss of differentiating cells at this late stage in embryonic development is a defect in progenitor maintenance resulting in the gradual depletion of progenitor cells. Evidence supporting this comes from a loss of radial glial fibres in the double morphant. Alternatively, without further analysis, it is possible that the double morphant embryos still contain progenitors but the onset of differentiation has been compromised and it is this that accounts for the phenotype. A model of the proposed function of Plzfa and Plzfb in progenitor maintenance in the 44 hpf hindbrain is shown in Figure 4-12.

A loss of progenitors is reminiscent of the described phenotype caused by knockdown of the Notch modulator *lunatic fringe* (*lfng*) (Nikolaou et al., 2009). The *lfng* knockdown embryos have significantly increased numbers of neurons caused by premature differentiation leading to less progenitor cells at later stages. I aimed to determine whether the double morphants displayed increased differentiation at an earlier stages of development, but I did not observe any significant increase in cells expressing *neurod4* or HuC/D at 30 hpf. This does not rule out premature differentiation completely, and may require analysis of other markers of neurogenesis, a different time point or a greater number of embryos to reveal any changes.

Figure 4-12: Model of Plzf function in the hindbrain

Cartoon depicting the proposed cause for the phenotype observed upon loss of both Plzfa and Plzfb function at 44 hpf. Injection of control MO results in hindbrain neurogenesis happening as described previously in Figure 4-7. Upon Plzfa and Plzfb knockdown there is a gradual depletion of neural progenitors such that by this stage there is a reduction of differentiating cells along the glial fibres. Consequently, neurons are no longer prevented from occupying locations adjacent to rhombomere boundaries in the mantle zone.

Figure 4-12


Control MO
Plzfa + Plzfb MO
44 hpf hindbrain (SV)
= ventricular zone
= mantle zone
= neural progenitor
= early differentiating cell (neurog1)
= late differentiating cell (neurod4)
= neuron

An alternative explanation is that the lower number of progenitors is caused by reduced proliferation. Technical reasons related to the developmental delay caused by the MOs hindered by ability to determine whether proliferation is affected in the morphants. Instead, I used gain of function experiments to see whether the number of cells undergoing mitosis was affected. I did not observe any significant changes in proliferation upon ectopic Plzfa expression in the early zebrafish embryo. Further analysis is required looking at other stages in the cell cycle in order to further investigate whether proliferation is affected by Plzf.

Comparisons with other model organisms

Studies in the mouse and chicken spinal cord found that Plzf is important for neural progenitor maintenance (Gaber et al., 2013). In these organisms there is just the single *Plzf* gene. I believe that I am observing a similar effect in the zebrafish hindbrain, although there are some differences when comparing results. They show that knockdown of protein function results in reduced expression of progenitor markers, such as Sox2. I carried out the same experiment in zebrafish, but did not observe any consistent changes in protein expression. Similar to my results, they do not observe an increase in terminal differentiation upon knockdown of Plzf in the chicken spinal cord. Instead, they see increases in early differentiation markers such as Neurog2, highlighting the requirement to test more markers in my analysis.

The differences in phenotypes between my work and the published work could be related to the different regions of the CNS being studied, as well as the different model organisms. They report that Plzf is gradually restricted to a central region of the dorsal-ventral axis in the spinal cord, and that Plzf positively regulates FGF receptor 3 in these cells (Gaber et al., 2013). This selectively provides cells with the ability to respond to the ubiquitously expressed FGF8 ligand, conferring a set of progenitors with greater proliferative potential than surrounding cells. Whether a similar mechanism occurs in the zebrafish hindbrain has not been directly tested. Evidence against the same mechanism occurring in zebrafish comes from the fact that I do not observe the same dorsal-ventral restriction of Plzf expression that they find.

Caveats to MO usage

The majority of my work into determining the role of Plzf within progenitors has involved knocking down expression using MOs. MO usage is commonly associated with off-target effects that can make interpretation of a specific phenotype difficult. Off-target effects include widespread cell death within the nervous system (Robu et al., 2007) and a developmental delay, as described in this work. It is therefore essential to include the correct controls to confirm the phenotype. This includes injecting MOs into a *tp53* mutant background where apoptosis is blocked and ensuring that the embryos are stage-matched rather than time-matched (Gerety and Wilkinson, 2011, Bedell et al., 2011).

There are specific problems relating to both using the *tp53* mutant line and the developmental delay. Blocking apoptosis is not ideal for investigating function of a gene that may include a role in cell survival. Several publications suggest that loss of function of Plzf results in increased apoptosis, including work in the CNS (Costoya et al., 2004, Parrado et al., 2004, Gaber et al., 2013). By preventing apoptosis for occurring in my MO experiments I may be not able to realise the full extent of the loss of function phenotype in zebrafish. Both my results and the published work suggest that attenuating Plzf function results in a subtle phenotype in the CNS (Gaber et al., 2013). The presence of a developmental delay therefore makes it extremely difficult to confirm that slight alterations in cell number are specifically related to Plzf function. Furthermore, MOs often do not lead to a complete loss of protein. I have evidence that the *plzfa* MO is able to significantly reduce protein expression from the antibody staining, but there may be low levels of protein remaining that are undetectable by the antibody.

Several methods have been used to confirm the specificity of knockdowns (Bill et al., 2009, Bedell et al., 2011). One method is to use several different MOs targeted against different regions of the same gene. RNA rescue is another important validation method. This involves providing both the MO and exogenous gene product into embryos and then determining whether this partially or fully rescues the phenotype. This technique is known to be problematic if there is a strong gain of expression phenotype for the gene, or overexpression affects development outside of the process being studied (Gerety and Wilkinson, 2011).

The definitive control for a MO knockdown is to observe the same phenotype in a mutant line for the same gene. Due to the uncertainties associated with MO usage, I aimed to generate loss-of-function mutations in the *plzfa* and *plzfb* genes and determine whether the resulting embryos display the same phenotype. This shall be described in detail in the next chapter.

Conclusions

In various tissues, Plzf has been shown to maintain progenitors and is important for the subsequent generation of differentiated cell types. In this chapter I have shown that *plzfa* and *plzfb* are widely expressed throughout the developing zebrafish CNS and overlap in neural progenitors of the hindbrain. I have presented data suggesting that Plzfa and Plzfb function together in the maintenance of progenitor cells. My next aim is to confirm the morphant phenotype and perform further analysis of Plzf function by generating zebrafish containing loss-of-function mutations in the *plzf* genes.

5 Editing the zebrafish genome using TALENs

5.1 Introduction

The zebrafish has several advantages for studying gene function, but has some limitations in the ability to carry out reverse genetic techniques to alter identified genes. Whilst MO usage has been widely adopted to transiently knockdown gene expression, they rarely lead to complete loss of protein (Bill et al., 2009), are unsuitable for disrupting multiple genes simultaneously and can lead to non-specific off-target effects, particularly in nervous system development (Robu et al., 2007, Gerety and Wilkinson, 2011). These off-target effects include a developmental delay, such as is shown in Appendix Figure 8-3, and the widespread induction of cell apoptosis, which can be rescued by the simultaneous knockdown of tumour suppressor protein Tp53 or the use of *tp53* mutants (Robu et al., 2007, Gerety and Wilkinson, 2011). Injection of the standard control MOs from Gene Tools does not appear to cause either of these toxic effects, making it difficult to accurately compare with the morphant embryos.

An ideal solution to these drawbacks would be to generate zebrafish lines containing a loss of function mutation in a targeted gene. There are ongoing large-scale attempts to generate mutant zebrafish lines using random mutagenesis followed by identification of affected genes (known as TILLING), but these efforts have not produced lines with mutations in the *plzf* paralogues at the time of writing (Zebrafish Mutation Project, Wellcome Trust Sanger Institute). In recent years, the advent of targetable nuclease technology has made it feasible to induce mutations at specific loci in the genome of numerous organisms, including zebrafish (Carroll, 2013). There are now three classes of genome targeting technologies: zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and the clustered, regularly interspaced, short palindromic repeats (CRISPR) Cas9 system.

I elected to use the TALEN technology with the intention of generating loss of function mutations in genes associated with *plzf*. The initial aim of generating mutants was to validate my knockdown experiments by determining whether I could phenocopy the morphant phenotype. Subsequently, I aimed to carry out further functional analysis on the targeted genes and perform a genome-wide screen for

expression levels of affected genes in the mutants in order to determine potential downstream targets.

In this chapter I will initially describe the strategy I used for inducing mutations in the zebrafish genome using TALENs and the methods used to screen for these mutations. I then show how I can efficiently generate loss of function mutations in both *plzf* paralogues and that these mutations are transmitted to future generations. Finally, I carry out functional analysis using these mutant embryos, attempting to phenocopy the morphant phenotype previously described in Chapter 4.4.

5.2 Targeting the genome using TALENs

ZFN and TALEN architecture

Current technology for genome editing employs natural DNA binding elements to generate double-stranded breaks in an organism's genome. Chronologically, the first technology developed for this purpose was ZFNs (Kim et al., 1996b). A ZFN is composed of a pair of arrays, each containing three or four DNA binding zinc fingers fused to the nuclease domain of the *FokI* type IIS restriction enzyme. In order for the enzyme to cut DNA, the *FokI* nuclease domain must dimerise (Bitinaite et al., 1998) and therefore double-stranded breaks are only induced when the pair of arrays binds DNA close to each other in the genome. Each zinc finger motif contacts three nucleotides and, in theory, multiple different fingers can be assembled together in order to target a specific DNA sequence. In practice, however, there are context-dependent effects whereby neighbouring zinc fingers influence DNA binding specificity and prevent this modular approach of assembly from being fully realised (Ramirez et al., 2008). Instead, labour-intensive and time-consuming selection-based methods are used in order to assemble these arrays (Foley et al., 2009).

TALENs recognise DNA in an apparently modular fashion, free from the context effects that hinder widespread use of ZFNs (Bogdanove and Voytas, 2011). Their DNA recognition domain (TALE) comes from certain plant pathogenic bacteria, in particular members of the *Xanthomonas* genus, which use TALEs to drive transcription of specific host genes (Römer et al., 2007). Each TALE protein contains

a conserved N-terminal domain, followed by a series of amino acid repeats and finally a conserved C-terminal domain. These natural TALEs have been reengineered to be suitable for use in a number of organisms and fused to the *FokI* nuclease domain to form artificial TALENs as shown in Figure 5-1a (Miller et al., 2011). It has been described that truncating the N- and C-terminus (described in the literature as Δ152 and +63) significantly improves the efficiency of TALEN cutting (Miller et al., 2011, Bedell et al., 2012, Mussolino et al., 2011). In an analogous way to ZFNs, TALENs work in pairs containing a left array and a right array.

The repeats found in nature are typically composed of 34 amino acids that are almost exclusively polymorphic within two amino acids at positions 12 and 13, referred to as the "repeat variable di-residue" (RVD), which confer base specificity (Moscou and Bogdanove, 2009, Boch et al., 2009). There are four RVDs most commonly found in nature, each recognising one of the four DNA nucleotides shown in Figure 5-1a. Three-dimensional structures of TALE-DNA complexes have been solved and reveal that amino acid 13 is responsible for contacting the base of the target nucleotide whilst amino acid 12 stabilises the structure (Deng et al., 2012, Mak et al., 2012). The structures also show that the NN and HD RVDs form hydrogen bonds with DNA bases whilst the other two RVDs make weaker van der Waals contacts.

In order to minimise the potential for off-target activity of the TALEN pairs, various groups have mutated specific residues within the cleavage domain of the *FokI* nuclease in order to alter its function from a homodimer to an obligate heterodimer (Miller et al., 2007, Szczepek et al., 2007, Doyon et al., 2011). The heterodimeric pair that I use contains mutations R487D (DD) and D483R (RR) for the left and right array respectively (Miller et al., 2007) (Figure 5-1b). These mutations eliminate off-target activity that could be caused by, for example, the left array binding in close proximity to another left array in the genome. It has also been reported that in some cases these heterodimeric TALEN pairs have a higher activity than their homodimeric counterparts (Cade et al., 2012).

Figure 5-1: Architecture of TALENs

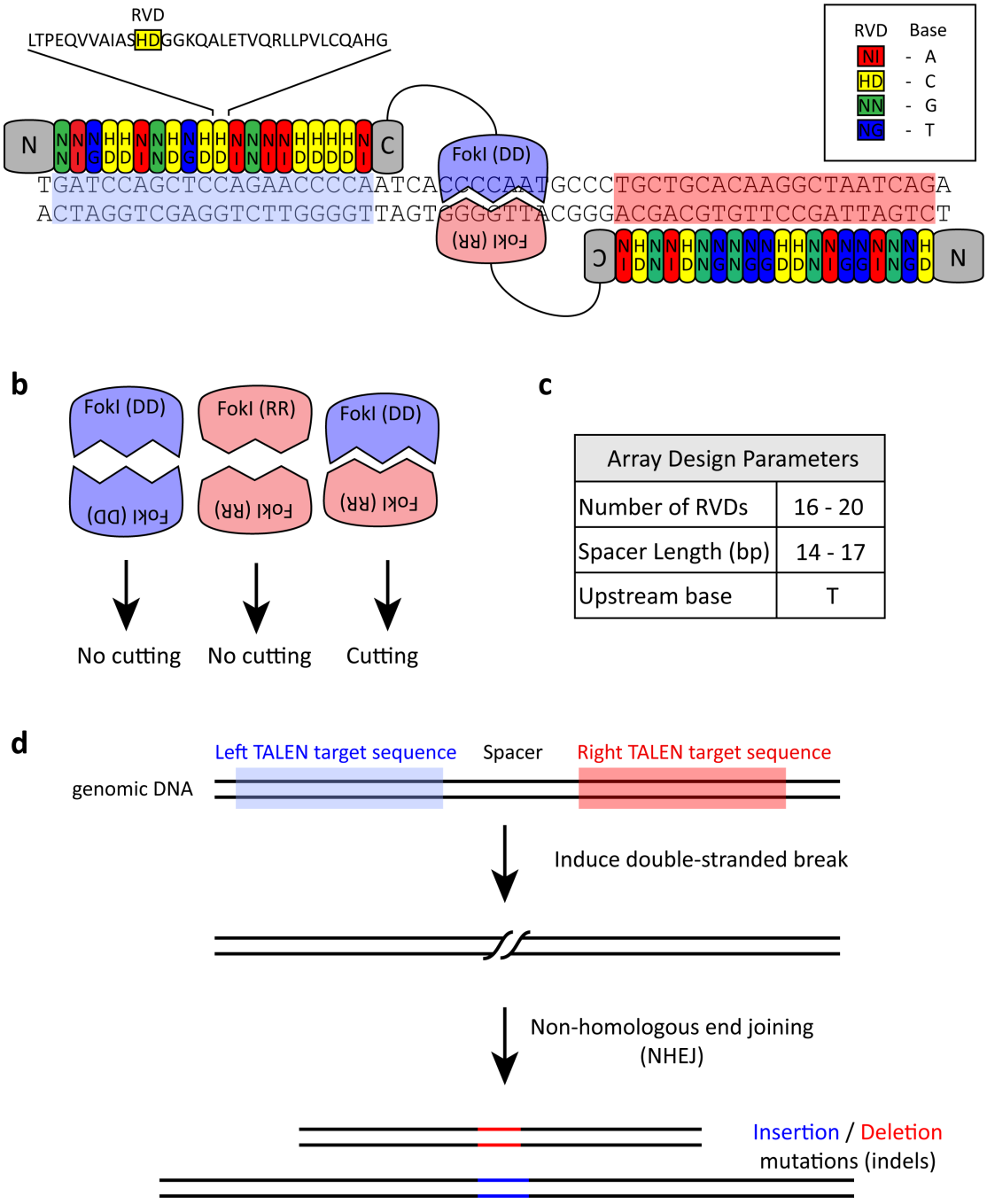
a: A heterodimeric TALEN pair. The left TALEN array recognises the 20 bp region highlighted in blue through the arrangement of RVDs shown whilst the right array recognises the 20 bp highlighted in red. The amino acid sequence for an individual repeat is shown and the RVD (HD) highlighted in yellow. The spacer is defined as the number of nucleotides between the left and right recognition sequences. RVDs are colour coded to show which base they interact with; NI, HD, NN and NG recognising A, C, G and T respectively (boxed). The recognition sequence shown is for the *plzfa* TALENs.

b: Heterodimeric TALENs. DNA cutting only occurs when the DD and RR *FokI* variants interact, which only happens when a left array and right array come into contact.

c: Parameters used for TALEN array design.

d: TALEN pairs induce double-stranded breaks in the genome that are erroneously repaired by the non-homologous end joining (NHEJ) pathway which frequently results in the insertion and/or deletion of nucleotides.

Figure 5-1


a
RVD
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
RVD Base
NI - A
HD - C
NN - G
NG - T
N
C
FokI (DD)
FokI (RR)
TGATCCAGCTCCAGAACCCCAATCACCCCAATGCCCTGCTGCACAAGGCTAATCAGA
ACTAGGTCGAGGTCTTGGGGTTAGTGGGGTTACGGGACGACGTGTTCCGATTAGTCT
b
FokI (DD)
FokI (RR)
FokI (DD)
No cutting
No cutting
Cutting
c
Array Design Parameters
Number of RVDs 16 - 20
Spacer Length (bp) 14 - 17
Upstream base T
d
genomic DNA
Left TALEN target sequence
Spacer
Right TALEN target sequence
Induce double-stranded break
Non-homologous end joining
(NHEJ)
Insertion / Deletion
mutations (indels)

Most natural TALEs have between 13 and 28 repeats (Bogdanove and Voytas, 2011) and there is evidence of a positive correlation with number of repeats and array activity (Cermak et al., 2011, Reyon et al., 2012). The spacer length is defined as the number of nucleotides between the sequences recognised by each array in a TALEN pair. The guidelines for the TALEN architecture I have elected to use are shown in Figure 5-1c (Dahlem et al., 2012). It has been reported that a thymine base just *5'* of the recognition sequence is required for high affinity binding (Mak et al., 2012, Lamb et al., 2013). Design and construction of the TALENs was carried out as described in Chapter 2.5. Synthesised capped mRNA encoding the TALEN pair was injected into the 1-cell stage of wild type zebrafish embryos.

Error-prone repair by non-homologous end joining (NHEJ)

There are two main pathways for the repair of DNA double-stranded breaks in eukaryotes: non-homologous end joining (NHEJ) and homologous recombination (Segal and Meckler, 2013). Homologous recombination repairs DNA breaks without error by copying information from a homologous sequence (West, 2003), described in more detail in Chapter 6. NHEJ is the predominant repair pathway for most organisms and is highly error-prone, often resulting in the insertion and / or deletion of nucleotides (known as indels) around the cut site.

The use of gene targeting technology to create loss-of-function mutations in target genes relies on our ability to exploit the natural NHEJ repair pathway. Targeting a pair of TALENs to the protein coding region of a gene causes a double-stranded break within the spacer region which when repaired by the NHEJ pathway will frequently cause missense and nonsense mutations that will lead to premature termination of protein translation, illustrated in Figure 5-1d. In order to maximise the potential effect of these mutations, I targeted the TALEN pairs close to the translational start site of the intended genes. In each case, I considered the possibility that a downstream in-frame ATG codon could result in a truncated, but still functional, protein even if the known translational start site is mutated. For *plzfa*, several ATGs were present within the first 50 bp of the coding sequence and therefore the TALENs for this gene were targeted just downstream of these.

Injected embryos are likely to be genetically mosaic for the target gene, containing a mixture of different mutant alleles as well as wild type copies. Numerous studies have provided evidence that TALENs are able to generate somatic mutations in zebrafish at efficiencies up to 100%, and that these mutations are transmitted through the germline (Huang et al., 2011, Sander et al., 2011, Dahlem et al., 2012, Bedell et al., 2012).

Methods of screening for mutations

TALENs are known to induce a range of mutations at the target site and therefore I sought a method to be able to reliably screen for the presence of these indels. For TALENs targeting the *golden* and *plzfa* locus I was able to screen for mutations by observing a phenotype. In the case of *golden*, mutations lead to a loss of pigmentation as previously described (Dahlem et al., 2012) and for *plzfa*, a loss of protein revealed by antibody staining. For genes where a clear phenotype could not be observed we turned to methods that detect mutations in the genomic DNA (gDNA) of the zebrafish.

Restriction fragment length polymorphism (RFLP) is a method that relies on a restriction enzyme recognition site within the spacer region that is likely to be lost upon formation of indels. This method has been successfully used to detect mutations caused by TALENs (Huang et al., 2011), but restricts the design of TALEN target sites to those containing a unique restriction enzyme site.

An alternative screening method that doesn't enforce restrictions upon TALEN design is High Resolution Melt (HRM) curve analysis, described in Figure 5-2. Genotyping by HRM relies on amplifying the target site in the presence of a double-stranded DNA intercalating fluorescent dye, followed by a slow denaturation of the DNA amplicon resulting in a decrease of overall fluorescence. Plotting this decrease against increasing temperature produces a curve similar to that shown in Figure 5-2b that contains information about nucleotide composition of the amplified DNA (Liew et al., 2004).

Comparisons of a wild-type curve and a homozygous curve will result in the same curve shape, shifted either to the left or right depending upon whether the mutation lowers or raises the melting temperature of the amplicon (compare black

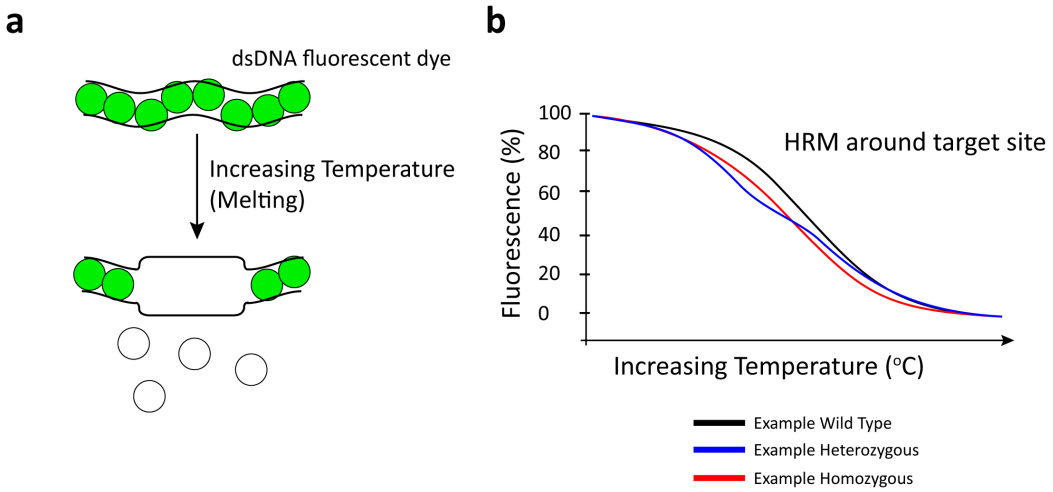
and red curves in Figure 5-2b). Heterozygotes produce more complex curves (blue curve in Figure 5-2b), which differ in shape to the wild type and homozygous curves as they are formed of multiple products containing different melt properties (Gundry et al., 2003). In genetically mosaic embryos, such as those injected with a TALEN pair, the melt curves become increasingly complex, but can easily be distinguished from wild type curves providing the mutation is present in more than 2% of the population (Dahlem et al., 2012). I elected to use the HRM analysis method for the

Figure 5-2: Screening for mutations using High Resolution Melt (HRM) curve analysis

a: Roughly 100 bp of gDNA around the TALEN cut site is amplified in the presence of a double-stranded DNA intercalating dye which is fluorescent when bound. Slowly heating the amplicon results in a loss of fluorescent as the DNA denatures and the dye disassociates.

b: Plotting normalised fluorescence as a percentage against increasing temperature results in melt curves that depend upon the nucleotide composition of the amplified gDNA. Comparing curves to the example curve representative of wild type DNA (black line): the homozygous curve (red line) is the same shape but is shifted as it has a lower melting temperature; the heterozygous curve (blue line) has a different shape.

Figure 5-2


a
dsDNA fluorescent dye
Increasing Temperature
(Melting)
b
HRM around target site
Fluorescence (%)
100
80
60
40
20
0
Increasing Temperature (°C)
Example Wild Type
Example Heterozygous
Example Homozygous

majority of my screening, using other methods to confirm the genotype when required.

Testing TALEN efficiency by targeting the golden gene

In order to test that I was able to induce mutations in target genes using TALEN technology and to assess the use of HRM analysis in screening for mutation I used a previously published TALEN pair targeted against the *golden* gene (Dahlem et al., 2012). The *golden* gene, also known as *slc24a5*, governs pigmentation of skin melanophores in zebrafish and homozygous mutants show a striking lack of pigmentation during early development (Lamason et al., 2005). This phenotype can most clearly be observed in the Retinal Pigmented Epithelium (RPE) surrounding the eye that is normally darkly pigmented in both the wild type and heterozygote at 48 hpf (Dahlem et al., 2012).

Figure 5-3a shows the experimental plan for the *golden* TALEN injections. Previous experiments reported that injected embryos display a high degree of genetic mosaicism (Dahlem et al., 2012), with some embryos displaying high numbers of mutant cells and others displaying fewer. I therefore elected to inject RNA encoding a fluorescent protein (H2B-RFP) alongside the *golden* TALEN RNA, hypothesising that those embryos containing high levels of RFP would contain the highest degree of mutations. Embryos were separated at 24 hpf into 'low RFP' and 'RFP Selected' pools and allowed to develop for a further 24 hours, at which point they were scored for the amount of pigmentation observed around the RPE (Figure 5-3b & c). Embryos with no RFP were assumed to have not been injected and were removed from the analysis. After a further day of development, individual embryos were lysed and HRM was carried out using previously described primers (Dahlem et al., 2012) (Figure 5-3d).

All of the uninjected embryos showed normal levels of pigment around the RPE whilst those injected with *golden* TALENs showed a range of pigmentation phenotypes, some of which are shown in Figure 5-3b. The presence of cells with no pigmentation suggests that the TALENs have induced mutations in both copies of the *golden* gene in those cells. The embryos were scored, ranging from wild type levels of pigment to no pigment and displayed in Figure 5-3c. I observed that embryos

from the RFP Selected pool had lower levels of pigmentation than those from the low RFP pool. This suggests that selecting for RFP correlates with TALEN activity in the injected embryos.

All the injected embryos screened by HRM analysis showed curves that significantly deviate from the wild type curve (Figure 5-3d). HRM curves that were produced from the RFP Selected embryos deflected further from the uninjected curve than the low RFP curves. This suggests that the further the HRM curve deviates from the wild type, the higher degree of mutations present in the target locus.

Germline transmission of mutations

Some of the RFP Selected injected embryos were raised to adulthood in order to test germline transmission of the mutations. Founder adults were initially outcrossed to wild type fish and their progeny were screened for a pigmentation phenotype at 48 hpf and by HRM at 72 hpf. All embryos were observed to have wild type levels of pigmentation (Figure 5-4a), indicating that they are either wild type or heterozygous for *golden*. Individual embryos from each founder were analysed for mutations in the *golden* locus by HRM and it was observed that all founders screened transmitted mutations to their progeny (8/8 founders). Figure 5-4b shows curves obtained from embryos from one of these founder fish.

All HRM curves observed were either indistinguishable from wild type curves or were reminiscent of the shape associated with heterozygous mutations shown previously in Figure 5-2b. The degree of transmission of mutations from individual founders ranged from 38% (3/8 embryos) to 100% (8/8). In clutches of embryos from founders I observed several different shaped curves (Figure 5-4b), indicating that a single founder can transmit multiple different mutant alleles in agreement with previous reports (Dahlem et al., 2012). Two individual founders were incrossed and some of the resulting progeny was observed to contain no pigmentation (Figure 5-4c) suggesting that these embryos are homozygous mutant for *golden*.

Figure 5-3: TALENs targeted against the *golden* gene generate embryos with a high number of biallelic mutations

a: Cartoon showing the experimental plan for assessing *golden* TALEN activity in zebrafish embryos. Embryos were injected with RNA encoding the TALEN pair (100 pg of each array) along with RNA for H2B-RFP (30 pg). At 24 hpf, embryos containing high levels of RFP were separated from those containing low levels. At 48 hpf, embryos were screened for the amount of pigmentation observed around the RPE. At 72 hpf, gDNA from the embryos was extracted and analysed by HRM.

b: Images showing representative embryos at 48 hpf. Wild type, uninjected embryos contained high levels of pigmentation around the RPE whilst those injected with *golden* TALENs show patches of cells without pigment.

c: 48 hpf embryos from the RFP Selected pool (n=33), the low RFP pool (n=23) and uninjected embryos (n=43) were scored based upon the amount of pigmentation observed around the RPE.

d: HRM curves generated from individual 72 hpf embryos. The RFP Selected (red) and low RFP (blue) curves deviate from the uninjected (black) curves. Each curve represents an individual replicate.

Figure 5-3

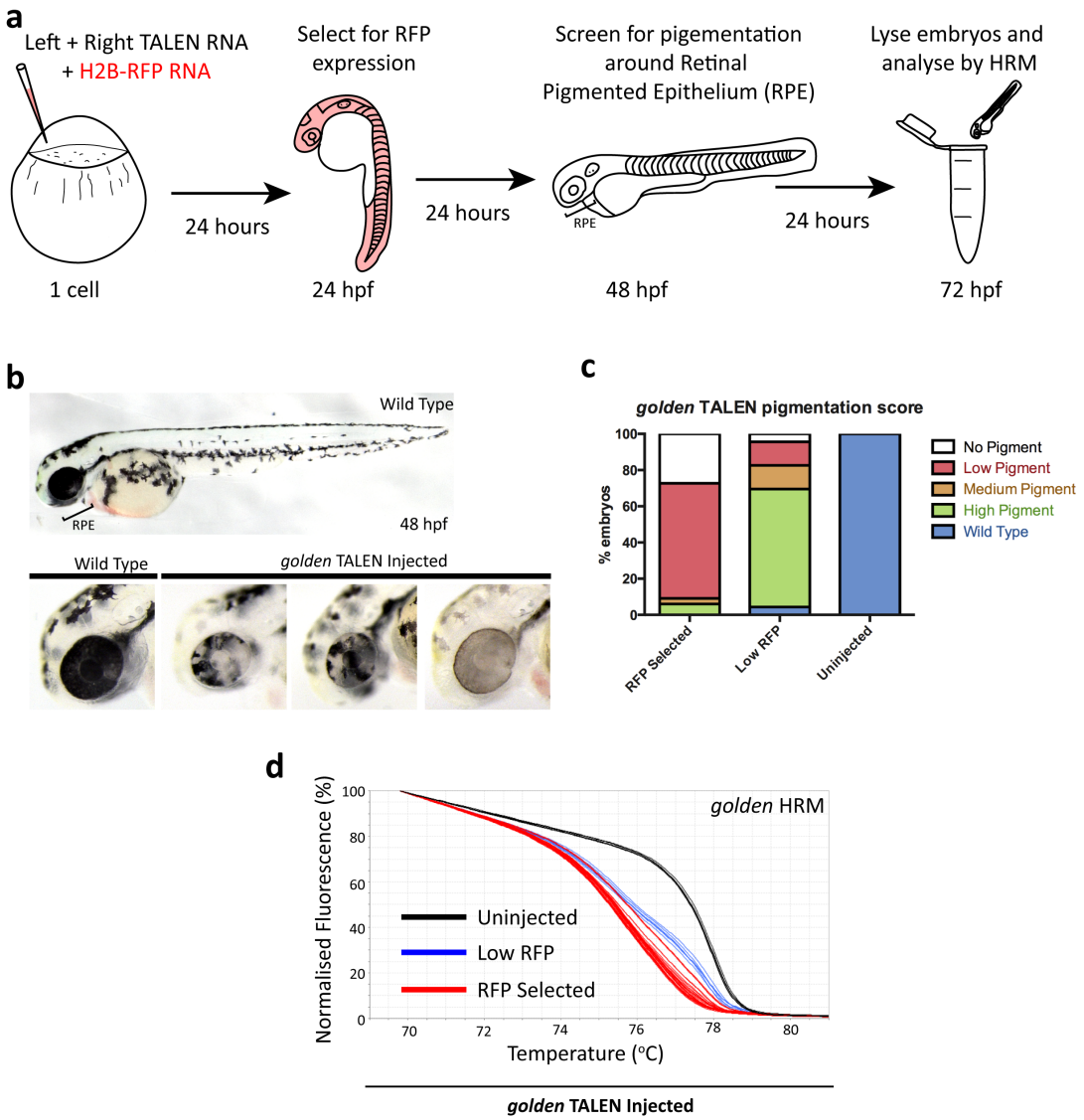

a
Left + Right TALEN RNA
+ H2B-RFP RNA
Select for RFP expression
Screen for pigementation around Retinal Pigmented Epithelium (RPE)
Lyse embryos and analyse by HRM
24 hours
24 hours
24 hours
RPE
1 cell
24 hpf
48 hpf
72 hpf
b
Wild Type
RPE
48 hpf
Wild Type
golden TALEN Injected
c
golden TALEN pigmentation score
% embryos
100
80
60
40
20
0
RFP Selected
Low RFP
Uninjected
No Pigment
Low Pigment
Medium Pigment
High Pigment
Wild Type
d
golden HRM
Normalised Fluorescence (%)
100
80
60
40
20
0
Uninjected
Low RFP
RFP Selected
70
72
74
76
78
80
Temperature (°C)
golden TALEN Injected

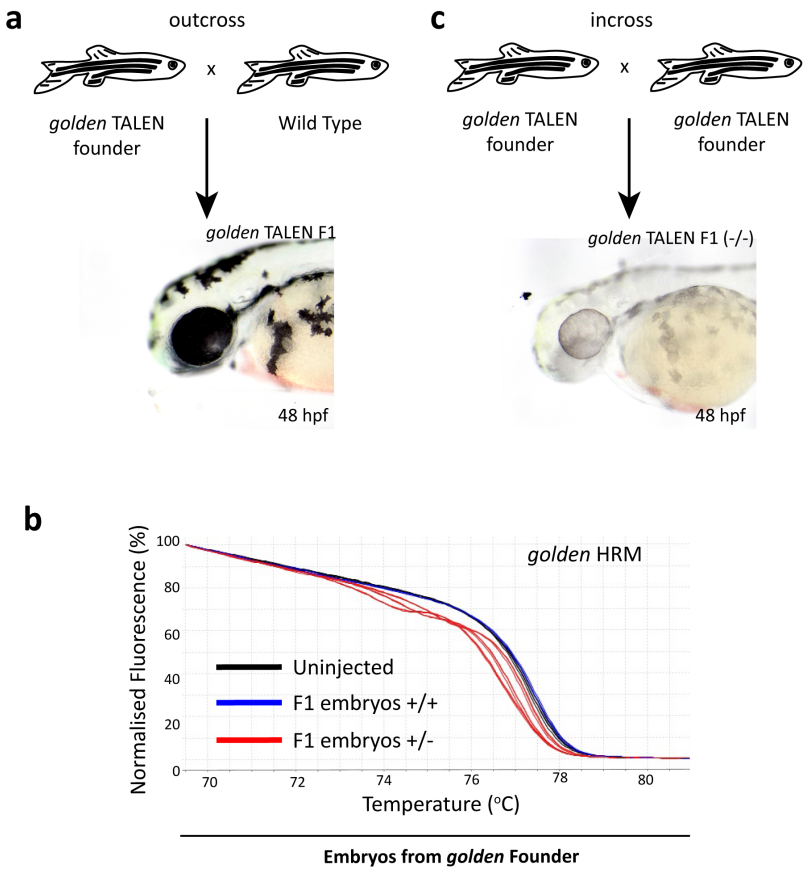
Figure 5-4: Transmission of the *golden* mutation through the germline

a: Adult *golden* TALEN founders were outcrossed to wild type fish and their progeny were examined to assess any pigmentation phenotype at 48 hpf. All embryos appear indistinguishable from wild type.

b: Individual embryos from clutches obtained from the outcrosses were analysed by HRM at 72 hpf. Shown here are HRM curves for several embryos from one clutch, demonstrating three different heterozygous curves (red) as well as curves (blue) that are indistinguishable from embryos obtained from wild type adults (black).

c: Adult *golden* TALEN founders were incrossed and their progeny examined. Amongst the clutches, some embryos appeared entirely pigmentless indicating that these embryos are homozygous mutant for *golden* (*golden* -/-).

Figure 5-4


a
outcross
x
golden TALEN founder
Wild Type
golden TALEN F1
48 hpf
c
incross
x
golden TALEN founder
golden TALEN founder
golden TALEN F1 (-/-)
48 hpf
b
golden HRM
Normalised Fluorescence (%)
100
80
60
40
20
0
70
72
74
76
78
80
Temperature (°C)
Uninjected
F1 embryos +/+
F1 embryos +/-
Embryos from *golden* Founder

Overall these experiments demonstrate that TALENs are suitable for efficiently generating loss of function mutations in target genes and that HRM analysis can be used to screen for mutations both in injected embryos and to assess germline transmission.

5.3 Generating *plzfa* and *plzfb* mutants

Using TALENs in target the plzfa locus

I designed and constructed a TALEN pair that would target a region of the zebrafish genome around 50 bp downstream of the start of the Plzfa protein coding region (Figure 5-5a). Wild type embryos were injected following a similar procedure as shown in Figure 5-3a, and embryos with high levels of RFP were analysed for evidence that mutations had occurred. Using the previously described zfPlzf Ab (Chapter 4.2) it is possible to monitor the effect injecting *plzfa* TALENs has on the endogenous Plzfa protein expression. Extensive loss of protein in up to 100% of cells in injected embryos was observed (Figure 5-5b-g) suggesting that the *plzfa* TALENs are highly efficient at inducing bialleic mutations.

The presence of somatic mutations at the *plzfa* locus was confirmed by carrying out HRM analysis on injected embryos, using primers designed around the target site (*plzfa* Target Site HRM). Melt profiles obtained from injected embryos deviate significantly from the uninjected profiles (Figure 5-5h). Finally, separating individual alleles from a single injected embryo and sequencing revealed a range of different deletions and insertions caused by NHEJ-mediated repair at the *plzfa* locus (Figure 5-5i).

plzfa TALEN specificity

As previously described, *plzfa* and its paralogues *plzfb* contain a high degree of sequence identity (Appendix Figure 8-1). It has been reported that TALENs are able to cut DNA at near homologous target sequences (Dahlem et al., 2012, Mussolino et al., 2011) and therefore it was important to determine whether the *plzfa* TALENs are able to also generate mutations at the *plzfb* locus. Comparisons of the two sequences reveal that the left TALEN array target site differs by two nucleotides, and the right TALEN array target site by four nucleotides (Figure 5-6a). Analysis by HRM at the *plzfb* locus (*plzfb* Off-Target Site HRM) homologous to the

plzfa Target Site resulted in curves indistinguishable from uninjected control curves (Figure 5-6a) suggesting that the *plzfa* TALENs are unable to induce mutations at the *plzfb* locus.

I examined whether I was able to titrate the efficiency of *plzfa* TALEN activity by injecting different amounts of RNA. Analysis by HRM revealed that observable mutations were induced when very little RNA was injected into the embryo (4 pg) and that increasing the amount injected correlated with an increasingly divergent melt profiles compared with the uninjected curve (Figure 5-6b).

***Targeting the plzfb* locus**

I used the same protocol to design and construct a TALEN pair to generate mutations in the protein coding region of the *plzfb* gene (Figure 5-7a). Analysis by HRM of the injected embryos indicated that these TALENs are able to induce mutations at the target site (Figure 5-7b). As described previously, I was unable to reliably detect the presence of the endogenous Plzfb protein in zebrafish using the zfPlzf Ab and therefore unable to use this technique to determine efficiency of the *plzfb* TALENs. The HRM curves didn't deviate as far from the wild type curves when compared with the *plzfa* TALENs (Figure 5-5h). Although HRM analysis doesn't allow quantification of the extent of mutations induced in individual embryos, I attribute this result to be indicating that *plzfb* TALENs aren't as efficient at inducing mutations compared to the *plzfa* TALENs. Increasing the amount of *plzfb* TALEN RNA injected didn't appear to have any significant effect upon the HRM curves produced (*data not shown*).

The high sequence identity between the *plzfb* TALEN target sequence and the homologous region in *plzfa* is shown in Figure 5-7c. As before, I observed no obvious alterations in the HRM curves from injected embryos when compared to the uninjected embryos (Figure 5-7d), indicating that *plzfb* TALENs are not generating indels in the *plzfa* gene.

Figure 5-5: TALENs targeting *plzfa* induce mutations with a high efficiency

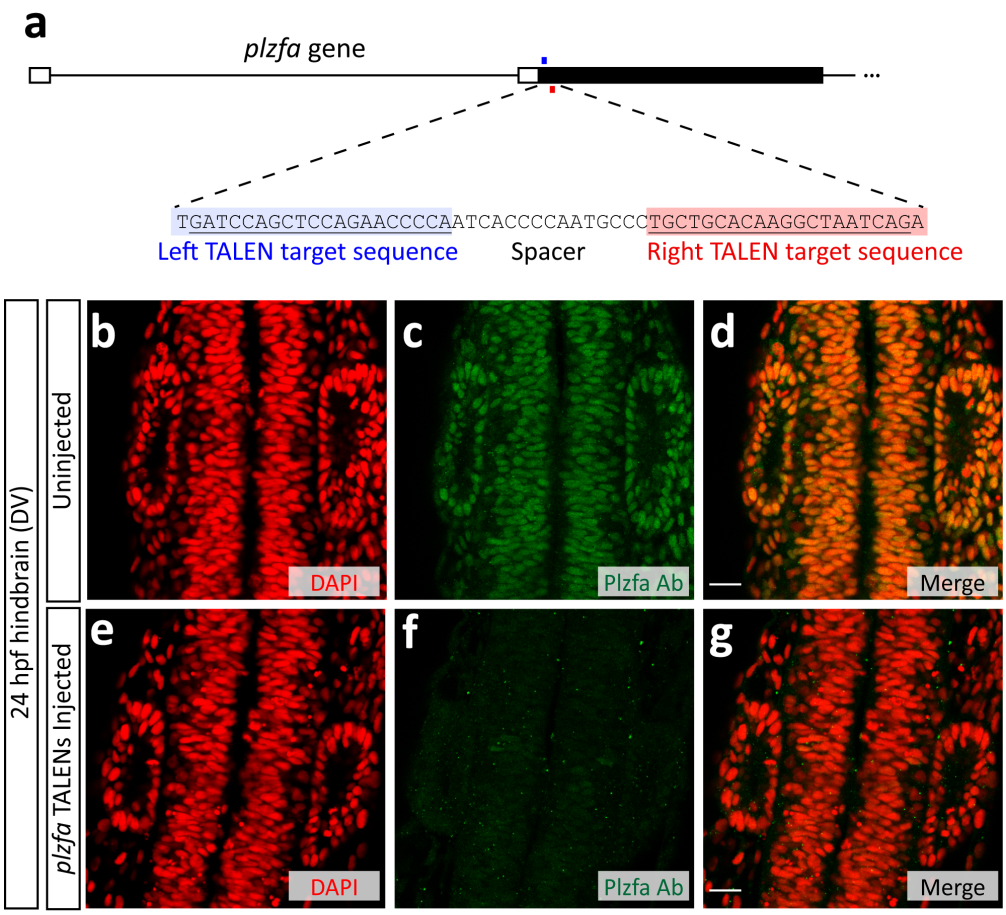
a: TALENs were designed to target the *plzfa* locus near the start of the protein coding region in the second exon. The forward DNA strand for the *plzfa* target sequence is shown, the sequence targeted by the Left *plzfa* TALEN highlighted in blue and the complementary sequence for the Right *plzfa* TALEN highlighted in red.

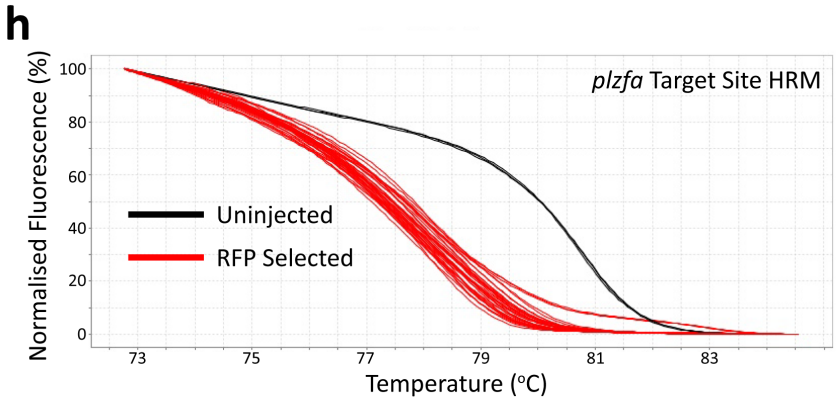
b-**g**: Using the zfPlzf Ab to monitor protein expression in embryos injected with *plzfa* TALENs RNA (100 pg each array). Uninjected embryos show widespread nuclear Plzfa protein expression in the 24 hpf hindbrain (**b**-**d**), which is entirely abolished in some embryos injected with *plzfa* TALENs (**e**-**g**). Scale bar = 20 µm.

h: Individual injected embryos were analysed by HRM analysis. All examined showed curves (red) that were noticeably different from the curves produced by uninjected embryos (black).

i: Various mutant *plzfa* alleles from a single injected embryo were separated by colony PCR and sequenced revealing a large range indels present in the injected embryo. Deletions are shown as a dash in red and insertions shown in blue.

Figure 5-5


a
plzfa gene
TGATCCAGCTCCAGAACCCCAATCACCCCAATGCCCTGCTGCACAAGGCTAATCAGA
Left TALEN target sequence
Spacer
Right TALEN target sequence
24 hpf hindbrain (DV)
Uninjected
plzfa TALENs Injected
b
c
d
e
f
g
DAPI
Plzfa Ab
Merge


h
plzfa Target Site HRM
Normalised Fluorescence (%)
Uninjected
RFP Selected
Temperature (°C)
plzfa TALEN Injected

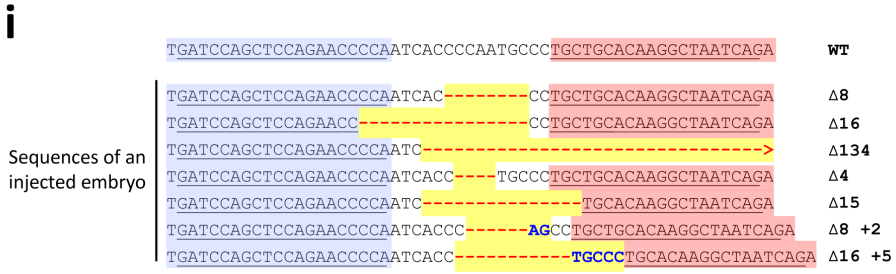

i
WT
Sequences of an injected embryo
Δ8
Δ16
Δ134
Δ4
Δ15
Δ8 +2
Δ16 +5

Figure 5-6: The *plzfa* TALENs do not target the *plzfb* locus and their efficiency can be titrated

a: Nucleotide sequence comparison between the targeted *plzfa* locus and the homologous region of the *plzfb* gene in the zebrafish genome. Left and right TALEN target sequences are underlined and an asterisk indicate the nucleotides not conserved between the two paralogues.

b: HRM analysis at the *plzfb* locus. The same embryos shown in Figure 5-5h were amplified around the *plzfb* locus and examined for evidence of changes in nucleotide composition. The HRM curves appeared highly similar to those obtained from uninjected embryos.

c: Titrating *plzfa* TALEN activity. Increasing amounts of *plzfa* TALEN RNA was injected into the zebrafish embryo and the embryos were analysed by HRM at the *plzfa* locus. Amounts indicated correspond to each array in the TALEN pair.

Figure 5-6

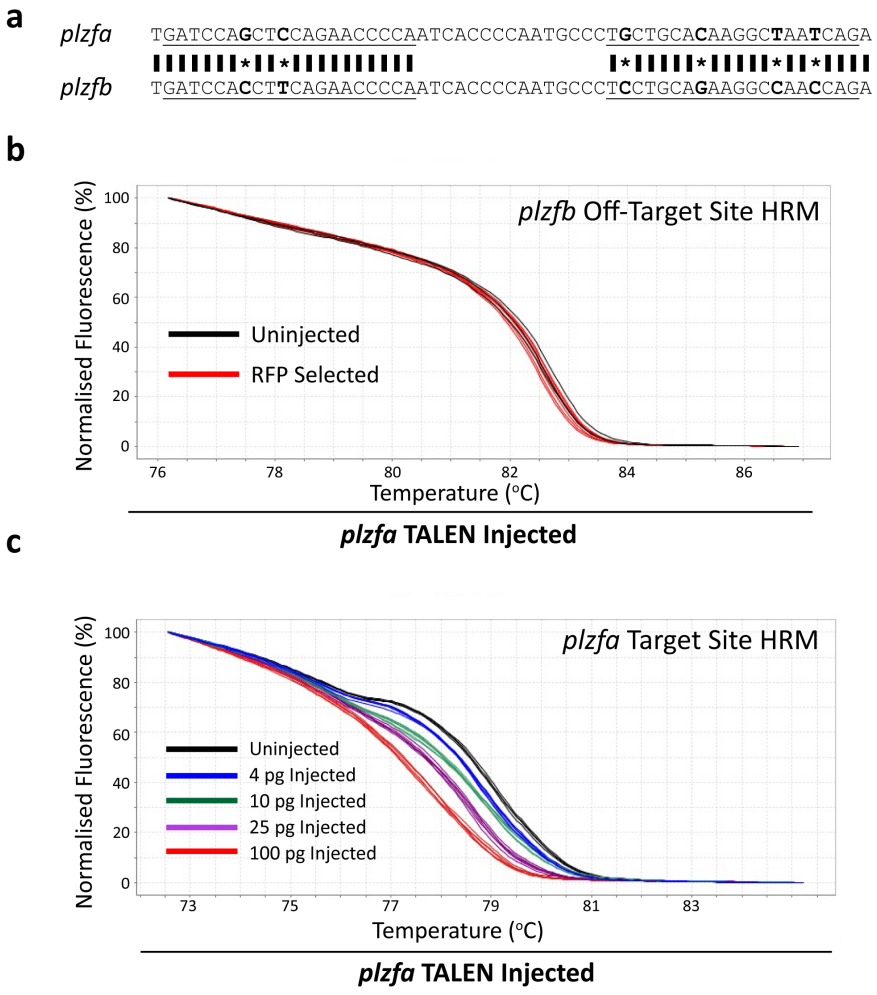
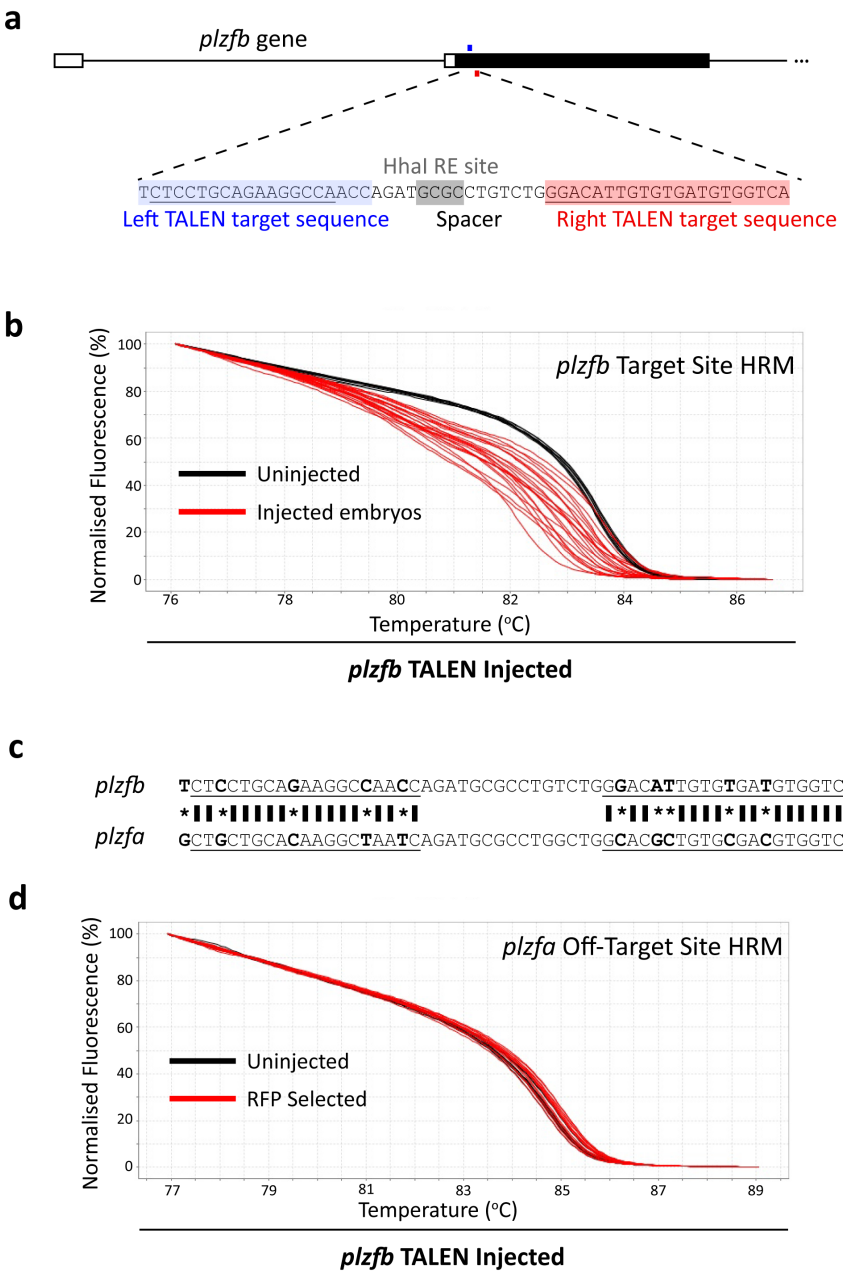

a
plzfa TGATCCAGCTCCAGAACCCCAATCACCCCAATGCCCTGCTGCACAAGGCTAATCAGA
plzfb TGATCCACCTTCAGAACCCCAATCACCCCAATGCCCTCCTGCAGAAGGCCAACCAGA
b
plzfb Off-Target Site HRM
Normalised Fluorescence (%)
Uninjected
RFP Selected
Temperature (°C)
plzfa TALEN Injected
c
plzfa Target Site HRM
Normalised Fluorescence (%)
Uninjected
4 pg Injected
10 pg Injected
25 pg Injected
100 pg Injected
Temperature (°C)
plzfa TALEN Injected

Figure 5-7: The *plzfb* TALENs induce somatic mutations and don't target *plzfa*

a: Schematic representation of the 5' *plzfb* gene in the zebrafish genome showing the first two exons. *plzfb* TALENs were targeted roughly 80 bp downstream of the translational start site. TALEN target sequence for the left array is highlighted in blue and the complementary sequence for the right array in red. The presence of the restriction enzyme site *HhaI* within the spacer region is highlighted in grey.
b: HRM analysis of *plzfb* TALEN injected embryos at the *plzfb* Target Site. The melt profiles for those embryos injected with *plzfb* TALENs deviate from the uninjected profile.
c: Sequence comparison between the *plzfb* target region and the homologous region in *plzfa*. Left and right target sequences are underlined and nucleotides which differ are marked by an asterisk.
d: HRM analysis for *plzfb* TALEN injected embryos at the *plzfa* Off-Target Site. No obvious differences from the uninjected melt profile were observed.

Figure 5-7


a
plzfb gene
HhaI RE site
TCTCCTGCAGAAGGCCAACCAGATGCGCCTGTCTGGGACATTGTGTGATGTGGTCA
Left TALEN target sequence
Spacer
Right TALEN target sequence
b
plzfb Target Site HRM
Uninjected
Injected embryos
Normalised Fluorescence (%)
Temperature (°C)
plzfb TALEN Injected
c
plzfb TCTCCTGCAGAAGGCCAACCAGATGCGCCTGTCTGGGACATTGTGTGATGTGGTCA
plzfa GCTGCTGCACAAGGCTAATCAGATGCGCCTGGCTGGCACGCTGTGCGACGTGGTCA
d
plzfa Off-Target Site HRM
Uninjected
RFP Selected
Normalised Fluorescence (%)
Temperature (°C)
plzfb TALEN Injected

Germline transmission of the plzfa mutations

Embryos injected with *plzfa* TALENs were grown up to adulthood. These fish were outcrossed to wild type fish in order to produce F1 embryos that are potentially heterozygous for *plzfa*. All embryos screened by HRM were shown to be heterozygous (24/24, *data not shown*). F1 embryos left to develop into adults were subsequently fin clipped and screened by HRM for the presence of indels around the target site (Figure 5-8a). All fish bar one were heterozygous (n=18) and, as expected from the previously described *golden* TALEN results, a range of mutations was found to be present in the F1 fish. These were clearly identifiable by their different melt profiles (Figure 5-8a). The mutant alleles of these fish were sequenced and found to contain indels within the spacer region (Figure 5-8b). Grouping the fish by their sequenced mutation gave the same result as grouping them by their melt profile. This result shows that it is possible to separate embryos carrying different mutations by their HRM curve shape.

plzfa TALEN founder fish were incrossed in order to generate F1 embryos that are potentially homozygous mutant for *plzfa*. Analysis by HRM revealed the presence of mutations in all embryos screened (n=12), and disruption of both *plzfa* alleles was confirmed by the absence of Plzfa protein revealed by antibody staining (Figure 5-8c-e). These results demonstrate that the *plzfa* TALENs are extremely efficient at inducing somatic mutations and that these mutations are transmitted to their progeny, allowing the generation of heterozygous and homozygous *plzfa* fish in a single generation. The results from the germline transmission of the mutations are summarised in Figure 5-8f. No obvious morphological phenotype or developmental delay was observed for the *plzfa* -/- mutants.

Germline transmission of the plzfb mutations

Generating *plzfb* mutants proved to be a more challenging task than for *plzfa*. *plzfb* TALEN injected fish were grown to adulthood and then incrossed in order to generate embryos with a mixture of genotypes. HRM analysis of the resulting embryos revealed a mixture of melt curves that appeared heterozygous (Figure 5-9a). From the HRM analysis alone, I was unable to determine whether the melt profiles were indicative of a heterozygous *plzfb* genotype, or whether they contained

two different mutant alleles as this would also result in melt curve that appeared heterozygous. Embryos from this founder were grown up to adulthood before further genotyping analysis was carried out.

I elected to use RFLP analysis on the F1 adults using the *HhaI* restriction enzyme site within the spacer region (Figure 5-7a). PCR was carried out around to amplify a 568 bp fragment around the TALEN recognition site. Analysis is complicated by the presence of a second *HhaI* site close to the one shown. If there are indels within the spacer region that disrupt the *HhaI* site then the loss of cutting can distinguish the genotype. The results, along with the expected fragment lengths for each of the genotype are shown in Figure 5-9b. Screening using this method found that out of the 15 fish screened, 8 were found to be heterozygous and the rest wild types with no homozygous mutants.

I determined the sequence of these indels and found several fish with the identical two base pair deletion (designated Δ2) shown in Figure 5-9c. This mutation leads to a premature stop codon at amino acid position 34 (total amino acids = 659) and is therefore is expected to lead to loss of Plzfb protein function. In order to generate mutant lines where both alleles contain this same mutation I crossed two of these identified adults (Figure 5-9d). A selection of the F2 embryos produced was screened by both RFLP (Figure 5-9e) and HRM analysis (Figure 5-9f) at 72 hpf, revealing a mixture of wild types, heterozygotes and homozygotes, as expected. The mutations were confirmed by sequencing and the chromatograms displayed in Figure 5-9g. No morphological phenotype was observed within the first five days of zebrafish development in any of the genotypes.

The remaining F2 embryos were grown to adulthood when they were fin clipped and analysed to identify their genotype by HRM (*data not shown*). The results from the germline transmission of the *plzfb* mutations are summarised in Figure 5-9h. These results taken as a whole suggest that the *plzfb* TALENs, although not as efficient at generating mutations when compared to the *plzfa* TALENs, are capable of generating transmissible mutations.

Figure 5-8: Highly efficient germline transmission of the *plzfa* mutations

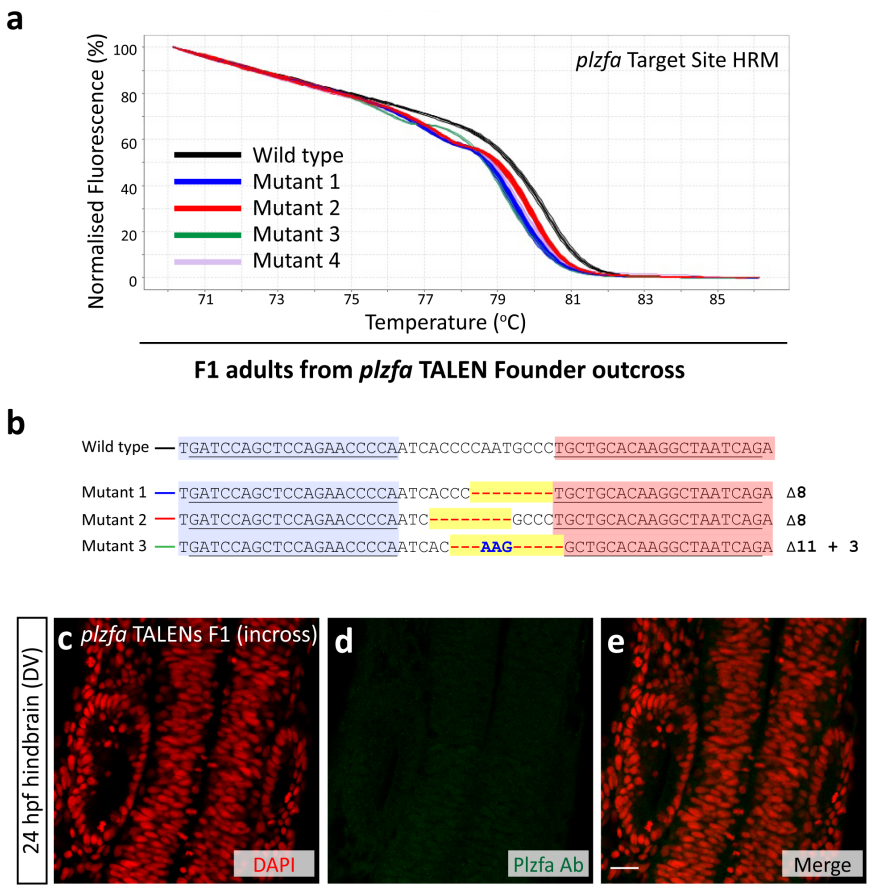
a: Melt profiles of *plzfa* TALEN F1 adults generated from a single outcross. Four different heterozygous curves were clearly distinguishable from the wild type curve. These are marked as mutant 1 (blue, 5/18), mutant 2 (red, 9/18), mutant 3 (green, 1/18) and mutant 4 (purple, 1/18). One fish was indistinguishable from wild type (not shown).

b: Sequencing of the different mutations in the F1 fish. Three of the mutant alleles are shown and are colour coded with their melt profiles shown in (a).

c-e: Adult *plzfa* TALEN founder fish were incrossed and their embryos analysed by IHC. Representative embryo showing loss of Plzfa protein in the hindbrain is shown. Compare with wild type levels of Plzfa protein expression in Figure 5-5b-d. Scale bar = 20 μm.

f: Summary of the data obtained by HRM from the germline experiments.

Figure 5-8



f

plzfa TALEN Germline Results	
Number of Founders transmit mutation / total screened	7/7
Number of F1 embryos with mutation / total screened	63/64

Figure 5-9: Germline transmission of the *plzfb* mutations

a: HRM analysis on embryos produced from incrossing two *plzfb* TALEN founder fish. Three example curves from individual embryos are shown in red, indicating that a range of *plzfb* mutations are transmitted through the germline.

b: RFLP analysis on fin clipped *plzfb* TALEN F1 adults produced from a single incross. PCR amplified a 568 bp fragment which was cut by the restriction enzyme *HhaI* resulting in a range of different fragments depending upon the genotype as indicated. All samples (15 total) were found to be either heterozygous (blue) or wild type (black).

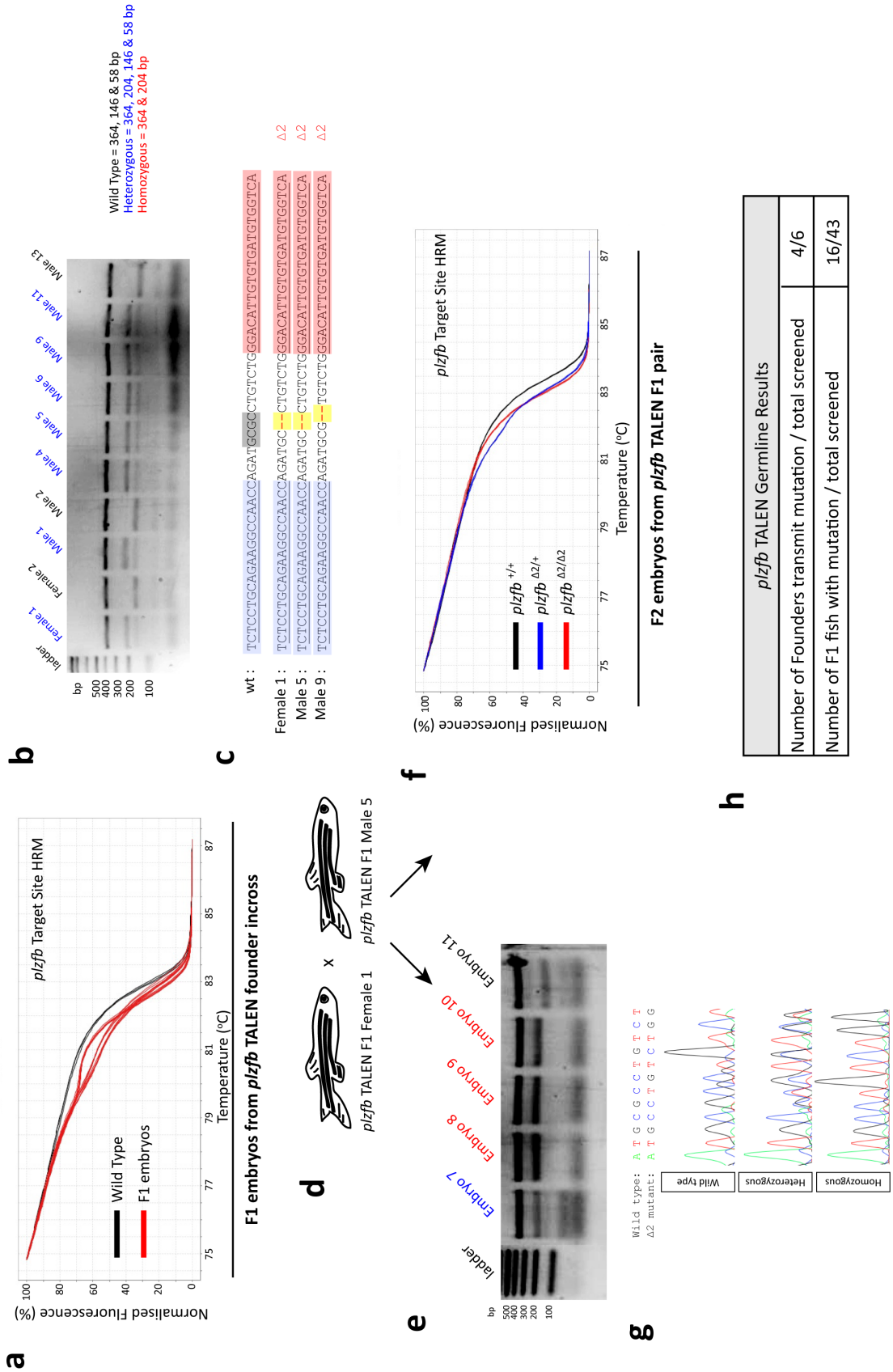
c: Sequencing some of the gDNA obtained from the F1 adults revealed a range of different mutations within the spacer. Female 1 and Male 5 were found to contain the same 2 bp deletion (Δ2) that leads to a premature stop codon in the *plzfb* gene. *HhaI* site highlighted in grey in the wild type sequence.

d-f: F1 Female 1 and Male 5 were crossed and their embryos analysed by RFLP (e) and HRM (f). Using either technique, three genotypes could be identified: wild type (black; 2/16 total fish), heterozygous (blue; 9/16) and homozygous (red; 5/16).

e: DNA chromatograms generated by the 4Peaks software shows the different profiles produced for wild types, heterozygotes and homozygotes. The expected sequences for the wild type and mutant alleles are displayed above the profiles, each nucleotide colour coded.

h: *plzfb* germline data summarised. Data pooled from both HRM and RFLP results.

Figure 5-9


a
Normalised Fluorescence (%)
plzfb Target Site HRM
Wild Type
F1 embryos
Temperature (°C)
F1 embryos from plzfb TALEN founder incross
b
Wild Type = 364, 146 & 58 bp
Heterozygous = 364, 204, 146 & 58 bp
Homozygous = 364 & 204 bp
bp
ladder
Female 1
Female 2
Male 1
Male 2
Male 4
Male 5
Male 6
Male 9
Male 11
Male 13
c
wt :
Female 1 :
Male 5 :
Male 9 :
Δ2
d
plzfb TALEN F1 Female 1
x
plzfb TALEN F1 Male 5
e
ladder
Embryo 7
Embryo 8
Embryo 9
Embryo 10
Embryo 11
f
Normalised Fluorescence (%)
plzfb Target Site HRM
plzfb +/+
plzfb Δ2/+
plzfb Δ2/Δ2
Temperature (°C)
F2 embryos from plzfb TALEN F1 pair
g
Wild type:
Δ2 mutant:
Wild type
Heterozygous
Homozygous
h
plzfb TALEN Germline Results
Number of Founders transmit mutation / total screened
4/6
Number of F1 fish with mutation / total screened
16/43

Generating double mutants

My MO experiments revealed a phenotype upon the knockdown of both *plzfa* and *plzfb* together. I therefore sought to generate a mutant zebrafish line carrying loss of function mutations in both genes. One method to do this is to cross *plzfa* $^{-/-}$ with *plzfb* $^{-/-}$ mutants, raise the progeny and then incross these fish to generate embryos with a mixture of genotypes, some of which will contain mutations in both paralogues. This method required waiting until I had genotyped adults of my single mutants before I could begin these crosses and was not suitable for the time scale of my PhD project.

Instead, whilst I was waiting for the single mutants to reach maturity, I attempted to generate double mutants by injecting RNA of both *plzfa* TALEN and *plzfb* TALEN pairs (100 pg each array) into wild type zebrafish embryos (Figure 5-10a). Over 90% of embryos injected with both TALEN pairs died in the initial 24 hours of development. The remaining survivors were screened by HRM analysis and determined to contain indels in the *plzfa* and *plzfb* genes (Figure 5-10b & c respectively).

It is unclear whether this toxicity relates to the increased amount of total TALEN RNA injected, is associated with increased off-target effects caused by injecting multiple left and right arrays, is a phenotype related to loss of *plzfa* and *plzfb* function or was due to injecting an unhealthy batch of embryos. The injections were repeated using half the amount of RNA (50 pg per array) and the vast majority of embryos were found to survive and looked morphologically healthy. HRM analysis on these embryos revealed that they contained identifiable indels in both genes (*data not shown*). Based on previous experiments, injecting lower amounts of RNA decreases TALEN efficiency in injected embryos and this is likely to decrease the efficiency of germline transmission of mutations.

I therefore elected to change strategy and inject the highly efficient *plzfa* TALENs into the *plzfb* mutant embryos in generate embryos with disrupted gene function in both paralogues. At the time the experiment was performed, I did not yet have the adult homozygous *plzfb* $^{\Delta 2/\Delta 2}$ mutants and I therefore injected the TALENs into the progeny of heterozygous F1 adults (Figure 5-10d). Analysis of the injected fish shows that, as expected, without injection of *plzfa* TALENs there are no

obvious indels in the *plzfa* locus (green line in Figure 5-10e) but upon injection of the TALENs indels are induced (red line in Figure 5-10e). The injected embryos did not display any obvious morphological defects during the first 5 days of development.

5.4 Using mutants to confirm the morphant phenotype

The loss of function results that I have obtained from MO experiments suggest that removing both *plzf* paralogues together results in a phenotype relating to a defect in progenitor maintenance. This phenotype is seen most clearly when examining the pattern of differentiated neurons in the zebrafish hindbrain at 44 hpf (Figure 4-8). In order to confirm the specificity of the MOs, I first wanted to discover whether I could observe the same phenotype with the TALEN-injected mutant embryos.

My experimental plan, shown in Figure 5-11a, was to cross the *plzfb* heterozygous fish described previously (Figure 5-9) to generate a mixture of wild type, heterozygous and homozygous *plzfb* embryos. I would then inject the *plzfa* TALENs into these embryos and allow them to develop until 44 hpf in order to analyse the phenotype by IHC. In order to determine the *plzfb* genotype of the individual embryos, the posterior part of the embryo was cut off after IHC and the gDNA extracted and analysed by HRM. This method allowed unambiguous assignment of genotype; representative melt profiles are shown in Figure 5-11b and a summary of all of the genotyped embryos shown in Figure 5-11c.

The anterior half of the embryos was flat mounted and the presence of endogenous Plzfa protein and arrangement of differentiated neurons examined. Representative images are shown in Figure 5-11d-i. The extent of Plzfa protein disruption varied between embryos but it was possible to find *plzfb* $^{\Delta 2/+}$ and *plzfb* $^{\Delta 2/\Delta 2}$ embryos that had near complete abolishment of zfPlzfa antibody staining (Figure 5-11f & h). In all genotypes, the pattern of differentiated neurons in the hindbrain was indistinguishable from wild type (Figure 5-11e, g, i & j) and in no cases could I observe a similar phenotype to that described by using MOs and shown in Figure 4-8. I also examined the presence of glial fibres in the embryos and saw no differences (*data not shown*).

Figure 5-10: Generating double mutant *plzfa* & *plzfb* zebrafish

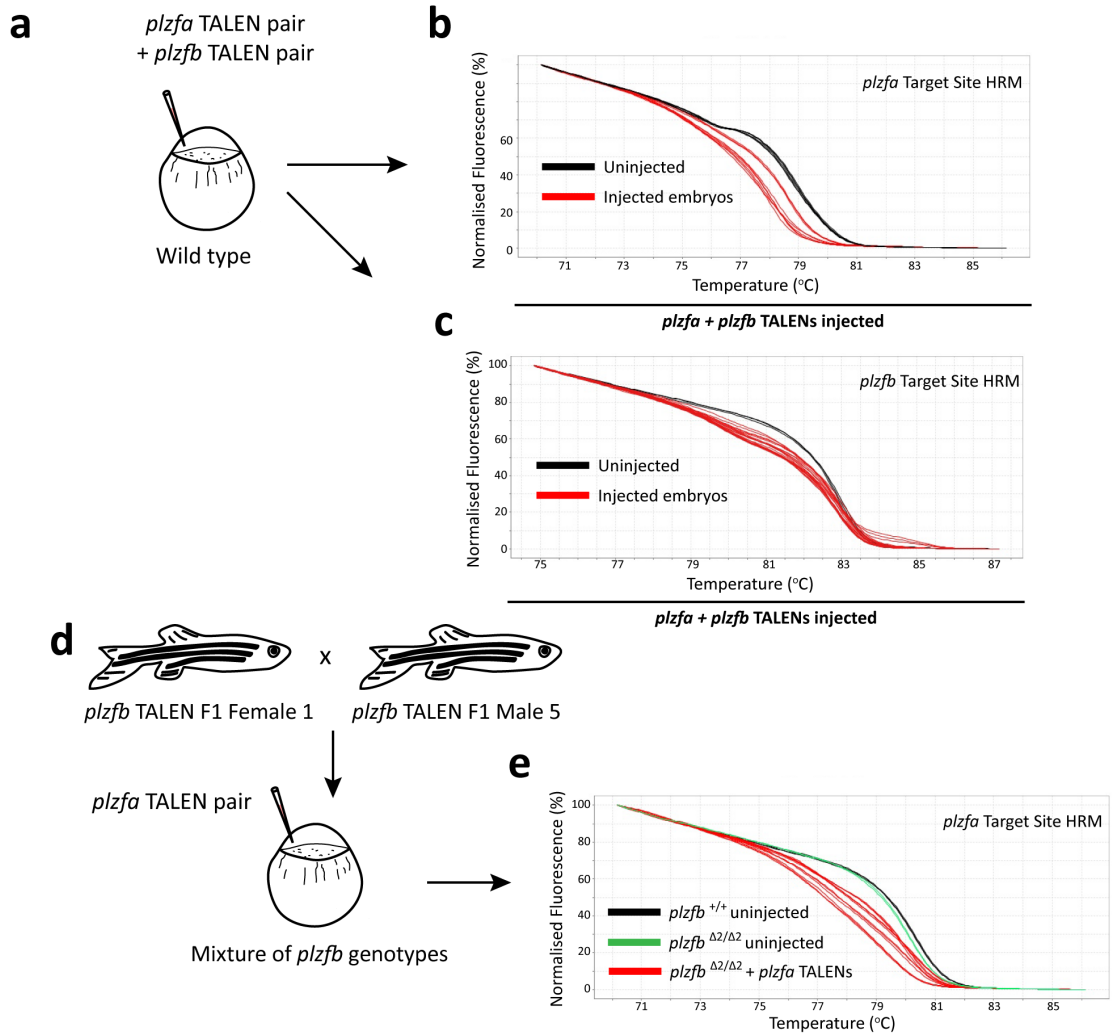
a: Wild type embryos were injected with RNA encoding the TALEN pairs targeting *plzfa* and *plzfb*. 50 pg of each array (400 pg total RNA) was injected.

b & **c**: The surviving embryos were analysed at 3 dpf by HRM analysis. Injected embryos were found to contain indels in the *plzfa* locus (b) and the *plzfb* locus (c). These results suggest that injecting the two pairs of TALENs can successfully induce mutations at both their target genes.

d: *plzfb* TALEN F1 heterozygotes previously identified (Figure 5-9) were crossed to produce a mixture of *plzfb* genotypes and *plzfa* TALENs were injected.

e: Melt profiles for *plzfb* $^{+/+}$ uninjected embryos (black), *plzfb* $^{\Delta 2/\Delta 2}$ uninjected embryos (green), and *plzfb* $^{\Delta 2/\Delta 2}$ embryos injected with *plzfa* TALENs (red). Deviations from the uninjected profiles can be observed in the injected profiles, indicating indels are present. *plzfb* genotypes were identified by HRM as in Figure 5-9f (*data not shown*).

Figure 5-10


a
plzfa TALEN pair
+ plzfb TALEN pair
Wild type
b
plzfa Target Site HRM
Normalised Fluorescence (%)
Uninjected
Injected embryos
Temperature (°C)
plzfa + plzfb TALENs injected
c
plzfb Target Site HRM
Normalised Fluorescence (%)
Uninjected
Injected embryos
Temperature (°C)
plzfa + plzfb TALENs injected
d
plzfb TALEN F1 Female 1
x
plzfb TALEN F1 Male 5
plzfa TALEN pair
Mixture of plzfb genotypes
e
plzfa Target Site HRM
Normalised Fluorescence (%)
plzfb +/+ uninjected
plzfb Δ2/Δ2 uninjected
plzfb Δ2/Δ2 + plzfa TALENs
Temperature (°C)

I carried out similar experiments by crossing *plzfa* $^{-/-}$ adult fish and injecting *plzfb* TALENs into the resulting embryos. As described, I have no clear method of assessing how much Plzfb protein has been affected in the injected embryos and therefore analysed embryos that had a high amount of H2B-RFP present (Figure 5-11k &m). Injecting either *plzfa* $^{+/-}$ or *plzfa* $^{-/-}$ with *plzfb* TALENs showed no disruption of differentiated neurons in the 44 hpf hindbrain (Figure 5-11l &n).

The above results show that the TALEN-injected embryos do not give the same phenotype observed when using MOs. A possible explanation for this result is that there is maternal RNA contribution of either of the *plzf* paralogues that, being unaffected by the injection of TALENs, could compensate for the loss of zygotic Plzf production. In order to account for this I crossed the adult fish that are *plzfa* homozygous mutant in order to generate embryos with neither maternal nor zygotic functional Plzfa protein. To remove any Plzfb function, I injected the *plzfb* MO that were used previously, together with *tp53* MO in order to block off-target effects caused by MO usage (Robu et al., 2007).

The experimental plan for doing this is shown in Figure 5-12a. Embryos were analysed for the presence of Plzfa protein and the organisation of differentiated neurons. Injecting the *plzfb* MO into either *plzfa* $^{+/-}$ or *plzfa* $^{-/-}$ embryos resulted in no obvious phenotype relating to a defect in progenitor maintenance at 44 hpf (Figure 5-12e & i respectively).

Figure 5-11: Functional analysis by injecting TALENs into mutant embryos

a: Experimental paradigm for analysing phenotype and genotype in mutant embryos. *plzfb* mutant genotypes were generated by crossing two heterozygous adults. The *plzfa* TALENs (100 pg each array) were injected into these embryos and at 44 hpf the embryos were fixed and kept separate for analysis. The anterior half of each embryo was separated and used to examine the phenotype by IHC (d-j) whilst the posterior was used to determine the genotype by gDNA extraction followed by HRM analysis (b & c).

b: HRM was carried out at the *plzfb* target site. Matching the melt profiles with previously genotypes samples identified wild types, heterozygotes and homozygotes. Representative samples are shown, along with the embryo label.

c: Summary of the *plzfb* genotypes obtained within the analysed clutch.

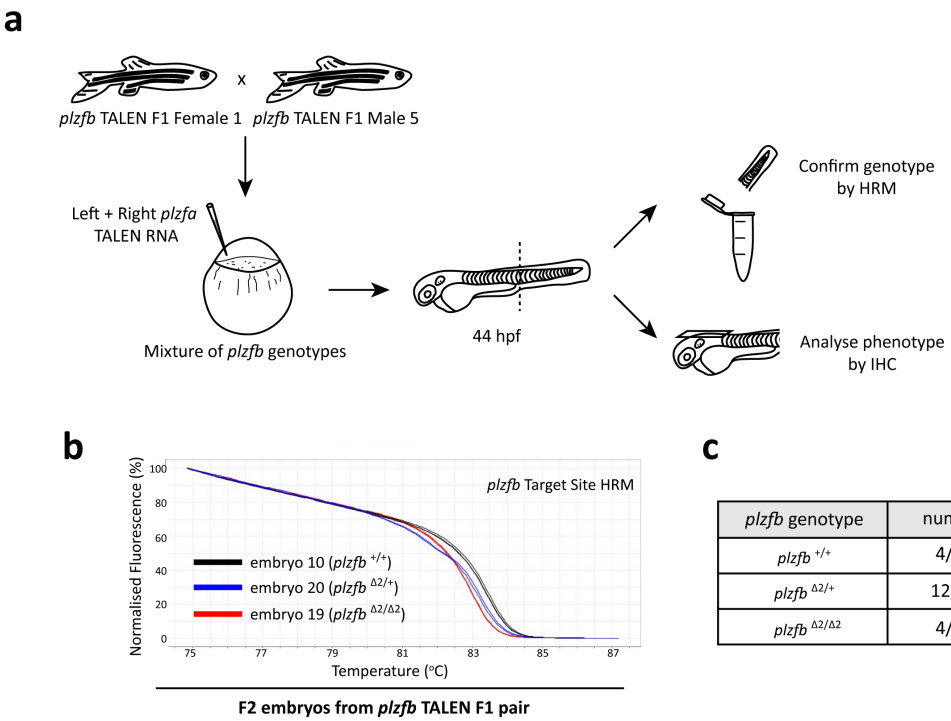
d-i: Phenotypic analysis of injected embryos by IHC. Dorsal views of the 44 hpf hindbrain of embryos from the 3 different genotypes are shown. In each embryo injection of the *plzfa* TALENs has significantly reduced the number of cells expressing Plzfa protein (d,f & h). The pattern of differentiated neurons (HuC/D) was indistinguishable from wild type in any of the examined embryos (e, g & i).

j: An example uninjected wild type embryo showing the stereotypical pattern of differentiated neurons in the 44 hpf zebrafish hindbrain.

k-n: *plzfa* $^{-/-}$ mutant fish were outcrossed to generate heterozygote embryos (k & l) or incrossed to generate homozygotes (m & n). 300 pg of each *plzfb* TALEN array was injected into these embryos along with H2B-RFP. Embryos with high numbers of RFP positive cells were analysed by IHC. In all cases, the pattern of differentiated neurons was unaffected (l & n).

Scale bar = 20 μm.

Figure 5-11



plzfb genotype	number
plzfb $^{+/+}$	4/20
plzfb $^{\Delta 2/+}$	12/20
plzfb $^{\Delta 2/\Delta 2}$	4/20

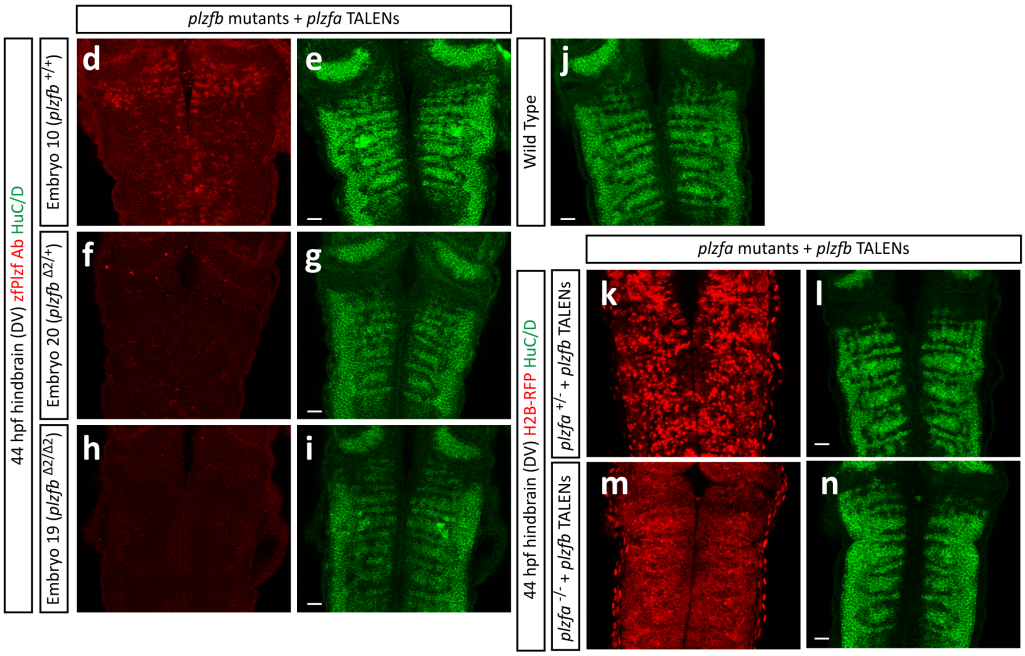


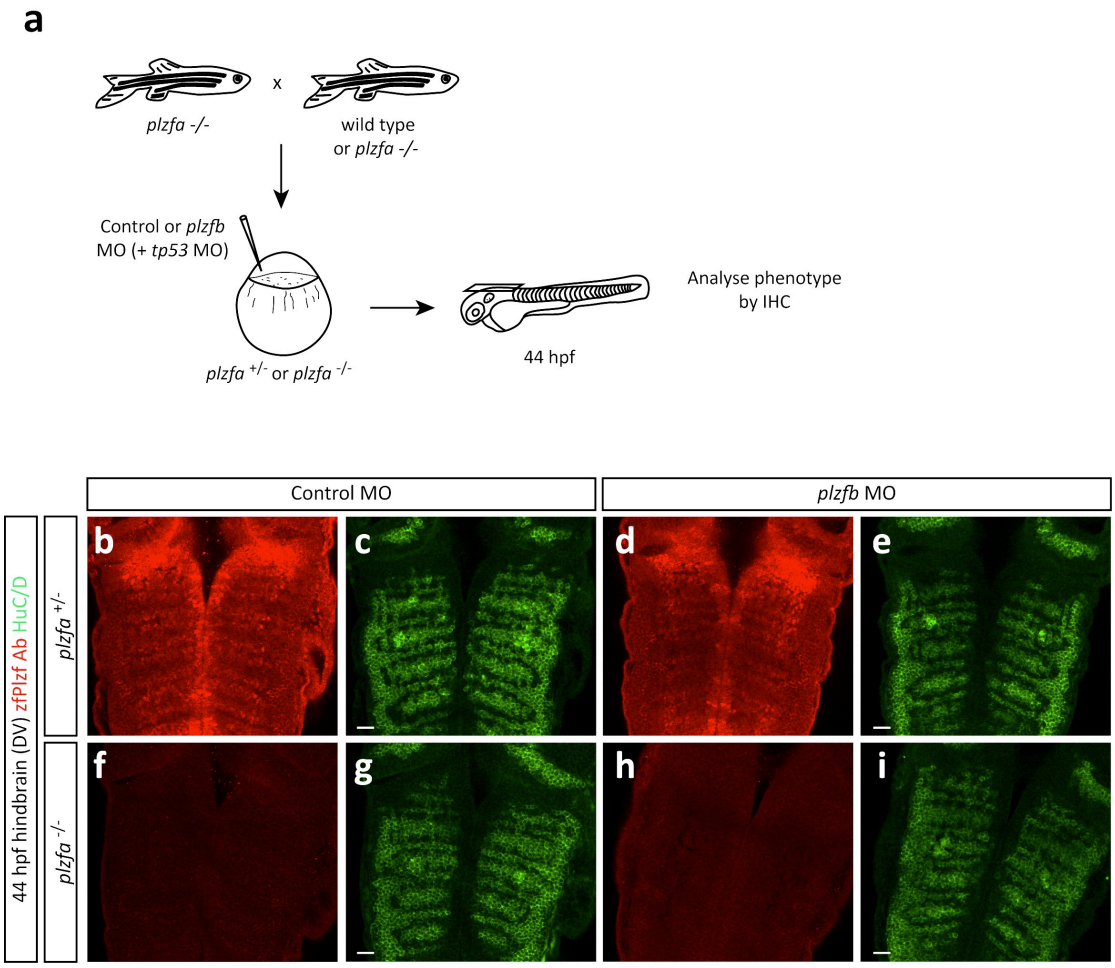
Figure 5-12: Functional analysis by injecting *plzfb* MOs into *plzfa* mutant embryos

a: Embryos heterozygous or homozygous for *plzfa* were produced as shown. 7.5 ng Control or *plzfb* MOs were injected into the embryos along with 7.5 ng *tp53* MOs and development proceeded as normal until 44 hpf when the embryos' phenotype was analysed by IHC.

b-i: Phenotypic analysis of MO injected mutants. Control MO or *plzfb* MO injected *plzfa* heterozygotes show wild type levels of Plzfa (b and d), whilst *plzfa* homozygotes show complete abloshment of Plzfa protein (f and h). No obvious phenotype was observed in any of the analysed embryos (c,e,g & i).

Scale bar = 20 µm.

Figure 5-12


a
plzfa -/-
x
wild type
or plzfa -/-
Control or plzfb
MO (+ tp53 MO)
plzfa +/- or plzfa -/-
44 hpf
Analyse phenotype
by IHC
Control MO
plzfb MO
44 hpf hindbrain (DV) zfPlzf Ab HuC/D
plzfa +/-
plzfa -/-
b
c
d
e
f
g
h
i

5.5 Discussion

TALEN efficiency, specificity and screening

The TALEN design parameters described by Dahlem et al. (2012) and shown in Figure 5-1c proved to be very effective at producing TALEN pairs that successfully target the genome. As well as the *plzfa* and *plzfb* TALENs described, I designed and injected TALENs for a further two genes associated with Plzf function. These are *redd1* and *btbd6a*, and HRM analysis showed the designed TALENs are able to induce mutations at their target locations (Appendix Figure 8-5). In total, I generated five different TALEN pairs and found evidence that four of them were able to induce indels at the target site, not including the previously published *golden* TALENs. The short turnaround from TALEN design to generation of capped mRNA ready to inject of just six days (Cermak et al., 2011, Dahlem et al., 2012) means that it was very simple to retarget the gene for which I was initially unsuccessful (*plzfb*, *data not shown*).

Whilst this project was underway, CRISPR / Cas9 technology for targeting genes was developed (Jinek et al., 2012, Hwang et al., 2013b). Rather than requiring large custom proteins to be built as is required for TALENs, CRISPR relies on the design of a single synthetic RNA molecule (guide RNA) to direct the Cas9 nuclease to its target site. This technique therefore has advantages over TALEN usage in the speed of construction and by using multiple different guide RNAs is capable of targeting several genes in the same cell using the same Cas9 nuclease (Cong et al., 2013). CRIPSPR technology has been shown to have similar efficiencies as TALENs at inducing mutations (Ding et al., 2013, Hwang et al., 2013b) (& unpublished data, Wilkinson lab) and although there has been some concern over their specificity (Fu et al., 2013), the technology is progressing at an incredible rate and improvements that address these concerns are being developed (Ran et al., 2013, Fu et al., 2014). The combined efficiency and ease of use of CRISPR suggests that it will, if it hasn't already, overtake TALENs as the method of choice for genome targeting.

My results have shown the benefits of injecting RNA encoding a fluorescent protein, in this case H2B-RFP, which can be used as an initial screening method for selected embryos with maximum TALEN activity. I build on previous data that

suggests HRM analysis is a simple and reliable method for screening for induced mutations (Dahlem et al., 2012). I show that it can be used to detect the presence of indels in the targeted locus of injected embryos and also to identify heterozygous and homozygous progeny. It is particularly useful once wild type, heterozygous and homozygous samples have been identified, as was the case for the *plzfb* F2 embryos. By including these controls into each analysis, the HRM software is able to identify which of the three genotypes each unknown sample belongs to without user input. Drawbacks of HRM are that it doesn't provide quantitative information regarding the number of cells that have been affected, nor can it tell us exactly what the nature of the indels are.

Having a second assay to validate my HRM screening results, such as phenotypic or RFLP analysis proved useful to confirm my genotyping. The genotyping process was more time-consuming that I expected, especially for the *plzfb* mutations that were transmitted to their progeny at less than 100% efficiency. As it seems likely that the majority of targetable nuclease technology will work at efficiencies less than 100%, these results show the importance of correctly genotyping. Had I spent more time to ensure that the *plzfb* F1 embryos tested in Figure 5-9a contained mutations in both alleles then I would have obtained homozygous mutant fish a generation (~ 3 months) sooner than I did.

It is not clear why the *plzfa* and *plzfb* TALENs function at different efficiencies (98% and 37% germline transmission respectively). As both TALENs are delivered into the embryos by the same method, the differences in efficiency likely arise from differences in the ability of the TALE DNA binding domains to target their respective DNA sequences. A potential explanation for why there are differences in efficiencies is that there are differences in accessibility of the DNA due to epigenetic modifications, which are known to affect TALE binding (Bultmann et al., 2012). Recent work has generated algorithms that predict which TALEN pairs will have a high activity (Lin et al., 2014). In some cases it is likely to be beneficial to decrease the efficiency of mutations induced, for instance if the gene knockout is embryonic lethal. The results for the *plzfa* TALENs (Figure 5-6c) demonstrate that TALEN efficiency can be titrated by altering the amount of RNA injected.

Previously published data has demonstrated that TALENs are able to induce mutations at off-target sites containing a high degree of homology to the target sequence (Dahlem et al., 2012, Guilinger et al., 2014). My analysis shows that neither the *plzfa* nor *plzfb* TALENs produce mutations detectable by HRM at their homologous off-target locus. A genome duplication event within the evolutionary history of the zebrafish (Postlethwait et al., 1998) means that many genes contain paralogues with high sequence identity and therefore it is important to confirm whether TALENs are able to target both paralogues in this organism. It remains unknown whether TALENs are able to induce other off-target effects that aren't related to sequence homology, such as is the case with MOs. Some care must therefore be taken when interpreting results from generated mutants before they have been outcrossed to wild type zebrafish for several generations. Generating embryos that carry two different mutated alleles for the same gene (transheterozygotes) could be useful for distinguishing off-target effects from effects relating to the loss of protein function.

Mutants versus MOs

The initial aim of generating mutations in the protein-coding region of *plzfa* and *plzfb* was to confirm the morphant phenotype described in the previous chapter. Unfortunately, I was unable to observe the same defect in progenitor maintenance in TALEN-injected embryos that I described for the double morphants. The fact that the two results don't correlate, suggests one of two possibilities. First, it is possible that the morphant phenotype is reflective of Plzf function and absence of phenotype in the mutant is because the mutations induced don't abolish protein function. The second possibility is that off-target effects, unrelated to Plzf function, cause the observed morphant phenotype.

Could the mutants still produce functional protein?

I shall begin by discussing the possibility that the mutants generated are not disrupting Plzf function. I targeted a region upstream of the BTB domain of the *plzfa* and *plzfb* genes with the TALENs in order to give me the maximum chance that any indels generated would ablate protein function. Making use of the antibody raised against the Plzfa BTB domain, I observe that the Plzfa mutants appear to eradicate

antibody staining. This result strongly suggests that the N-terminal BTB domain is not correctly produced. The importance of the BTB domain in driving Plzf nuclear localisation and carrying out transcriptional repression has previously been well documented, described in Chapter 1. Although I have been unable to use the antibody to confirm that the *plzfb* mutations also disrupts protein function, mapping the mutation to a 2 bp deletion that results in a premature stop codon close to the start of the BTB domain, suggests that this mutation would also lead to loss of function.

It is worth considering that the alternative splice products observed in the human genome generate truncated Plzf proteins that lack the BTB domain (Zhang et al., 1999). Although no function has been described for these proteins, if they exist in the zebrafish then the mutations I generated would have no effect on their function. These splice variants contain the same *5'* UTR as the full-length transcript and therefore would still be affected by the MOs. These alternative spliced products are not believed to be present in the mouse (Suliman et al., 2012) and analysis of the publically available RNA-Sequencing data suggests only the single transcript for *plzfa* and *plzfb* exist in the zebrafish genome (Collins et al., 2012). It therefore seems unlikely that this could account for the absence of phenotype in the mutants.

A Plzf isoform lacking the BTB domain has been observed to have a functional role in human colorectal cancer cell lines (Jones et al., 2013). This protein was produced from a second translational start site found downstream of the BTB domain and is associated with the cytoskeleton where it has a role in increasing cell adhesion and survival. I have not checked whether this shorter protein is expressed during zebrafish development, but the MOs against full length *plzfa* and *plzfb* are unlikely to affect translation of this protein. It therefore seems unlikely that this truncated protein could account for the discrepancy been the morphant and mutant phenotype.

There are two mouse mutants for *Plzf* that display phenotypes related to spermatogenesis and skeletal patterning (Barna et al., 2000, Buaas et al., 2004). The *luxoid* mutation results in a premature stop codon at 234 in the *Plzf* gene, in between the regions coding the BTB domain and the zinc fingers (Buaas et al., 2004). This mutation is downstream of the internal translation start site described by Jones

et al. (2013) and therefore would disrupt protein function of both the full-length murine Plzf and the truncated protein. A phenotype related to neural progenitor maintenance in the spinal cord was observed in the *luxoid* mice (Gaber et al., 2013). The authors describe that the same phenotype observed in the chicken spinal cord is dependent upon the repressive activity of Plzf, therefore suggesting that it is the full-length Plzf protein that is responsible for progenitor maintenance.

Removing the entire first coding exon of *Plzf*, responsible for coding over half of the protein, generated the second mouse mutant (Barna et al., 2000). This approach of removing large regions of gDNA results in a much higher chance that the mutation will result in a loss of function. The advent of multiplexing CRISPR technology has meant that we now have the ability to obtain these types of mutants in zebrafish (Cong et al., 2013, Ran et al., 2013). In hindsight, it may have been suitable to target similar regions to these mouse mutants as these are known abolish Plzf function and provide a phenotype.

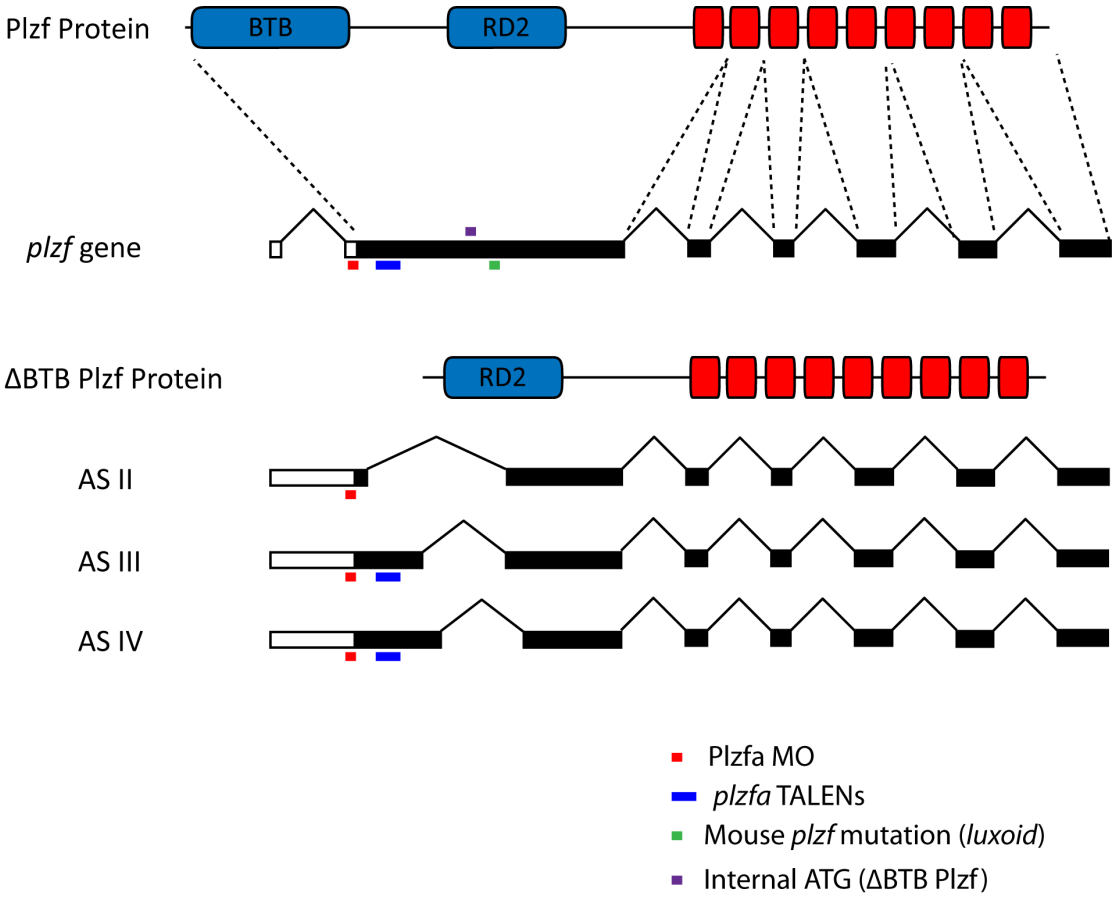
Finally, the possibility of maternally contributed Plzf protein was explored. Because of time constraints, I was unable to carry out functional analysis on embryos generated from double mutant adults. This means that there was the prospect that maternally contributed RNA or protein could be compensating for the loss of the zygotic protein. I tested this by injecting Plzfb MOs, which would affect both the maternal and zygotic transcript into embryos derived from *plzfa* mutant adults. The absence of a progenitor phenotype in these embryos suggests that maternal contribution is not compensatory.

Overall, I have evidence that the BTB domain is being disrupted in the *plzfa* and likely the *plzfb* mutants. This suggests that if any remaining protein is produced, it will lack the ability to multimerise, form transcriptionally repressive complexes and localise to the nucleus. I considered the existence of alternatively splice transcripts, protein isoforms and maternal contribution, but none could reasonably explain why I observe a phenotype in the morphants and not in the mutants. A schematic of the some of these *plzf* splice variants and both the full-length and truncated protein are shown in Figure 5-13, along with positions of the MOs, TALEN target and *luxoid* mutation.

Figure 5-13: Comparison of splice variants and isoform with the zebrafish and mouse mutants

Schematic of the *plzf* gene and protein, representative of either the human, mouse or zebrafish homologues. Location of the *plzfa* MO (red) and *plzfa* TALEN (blue) target regions are indicated. The internal ATG (purple) results in the production of the displayed truncated Plzf protein lacking the BTB domain (ΔBTB Plzf) in human cancer cells (Jones et al., 2013). This ATG is conserved in the zebrafish *plzfa* and *plzfb* genes, although whether the shorter protein is produced is unknown. Also indicated is the location of the luxoid mutation (Arg234 → Stop) from Buaas et al. (2004). Information on alternative splicing transcripts (ASII, III and IV) was obtained from Zhang et al. (1999).

Figure 5-13


Plzf Protein
BTB
RD2
plzf gene
ΔBTB Plzf Protein
RD2
AS II
AS III
AS IV
Plzfa MO
plzfa TALENs
Mouse *plzf* mutation (*luxoid*)
Internal ATG (ΔBTB Plzf)

Is the phenotype caused by off-target effects associated with MOs?

The association of MOs with off-target effects has been well documented. Toxic effects caused by some MOs include widespread cell death throughout the zebrafish CNS (Robu et al., 2007). Further research revealed that MO-mediated cell death leads to a decrease in proneural and neuronal marker expression in the CNS (Gerety and Wilkinson, 2011). This study also showed that pro-apoptotic genes are involved in the endogenous process of hindbrain boundary formation, independent of their role in cell death. Therefore an artifact of toxic MO usage is ectopic expression of these boundary markers (Amoyel et al., 2005, Gerety and Wilkinson, 2011). Both these effects can be rescued by simultaneous removal of Tp53, a tumour suppressor protein responsible for cell apoptosis (Robu et al., 2007, Gerety and Wilkinson, 2011).

It has been estimated that between 15-20% of MOs induce these toxic effects (Bedell et al., 2011). Unpublished work from the Wilkinson lab has shown that the *plzfa* MO causes widespread apoptosis and hindbrain boundary spreading and that this is rescued by removing Tp53 at the same time. There is also evidence that MO-mediated knockdown of Tp53 is unable to completely rescue the toxic phenotype and that injection into *tp53* mutants is encouraged (Wilkinson lab, unpublished results). Therefore the vast majority of my MO injections were done in the *tp53* mutant line.

The phenotype that I observe upon simultaneous injection of both *plzfa* and *plzfb* MOs is attributed to a defect in progenitor maintenance that is not reminiscent of the previously described MO-mediated toxicity. Nevertheless, there is a growing concern in the zebrafish community that there are some off-target effects of MOs that are independent of Tp53 activity. The widespread adoption of TALEN and CRISPR technology has meant that researchers are now starting to directly ascertain whether MO phenotypes occur in mutants.

An interesting recent publication tests exactly this (Law and Sargent, 2014). The authors previously described a phenotype related to defects in myelopoiesis upon injection of *pak4* MO (Law and Sargent, 2013). Coinjection of the MO with *tp53* MO results in the same phenotype and they are able to perform rescue

experiments by injecting mRNA encoding the wild type protein, suggesting the phenotype is specific to Pak4 function. However, analysis of *pak4* mutant zebrafish generated by TALEN technology does not phenocopy the morphant phenotype (Law and Sargent, 2014). The authors examine the possibility of splice variants and internal translational start sites that could account for this discrepancy, but find no evidence that this is the case. Instead, they conclude that Pak4 is dispensable in zebrafish and propose that the MO phenotype is caused by off-target effects unrelated to cell apoptosis. This study demonstrates that even when combined with robust control experiments, MO usage can result in misleading phenotypes.

I therefore propose that the MO-mediated phenotype that I describe in the previous chapter is not specific to the function of the Plzf proteins and instead is related to Tp53-independent off-target toxicity. Without further analysis it is difficult to determine exactly what is causing this phenotype, although possible causes will be discussed in Chapter 7.

Finally, it is worth considering whether the previously described phenotypes in the mice mutant for *Plzf* are present in the zebrafish models. The mice display a progressive loss of mature sperm and therefore the males are sterile as adults (Buaas et al., 2004). Plzf is expressed within spermatogonia of the adult zebrafish testis (Ozaki et al., 2011), however whether it has a homologous role to Plzf in the mouse has not yet been established. The *plzfa* mutant male zebrafish produce progeny, indicating that they are not sterile, however this has not been tested in the double mutant. The mice also display defects in hind limb patterning (Barna et al., 2000). Although a detailed analysis has not been performed, the homologous structure in the fish, the ventral fins (Grandel and Schulte-Merker, 1998), does not display any gross morphological differences in the adult *plzfa* mutant.

The *Plzf* mouse mutant showed increased production of certain neuronal cell types within the spinal cord (Gaber et al., 2013). The analysis I have carried out on the *plzfa* mutant zebrafish embryos only examined the pan-neuronal marker HuC/D. Therefore analysis with markers specific for neuronal subtypes is required to determine whether there is a neurogenic phenotype within the mutant embryos.

Conclusions

In conclusion, I have demonstrated that targetable nucleases, such as TALENs, can be used to induce mutations in genes associated with Plzf activity in the zebrafish. I present evidence of TALENs efficiency and specificity at inducing indels at the target genes and make use of several screening methods. Functional analysis of embryos containing mutations in both *plzfa* and *plzfb* did not lead to the same phenotype as the double morphants. I consider several possibilities that could account for this and propose that the morphant phenotype is a MO artifact caused by Tp53-independent pathways.

6 Targeted insertions into the zebrafish genome using TALENs

6.1 Introduction

Zebrafish embryos have a number of advantages that make them suitable for *in vivo* imaging, such as being largely transparent and developing outside the mother. The development of Tol2 transgenesis (Kawakami et al., 2004, Balciunas et al., 2006) as a technique to efficiently insert exogenous DNA into the zebrafish genome greatly increased the ability to analyse and make use of endogenous enhancers and promoters. The random integration at multiple locations in the zebrafish genome makes Tol2 transgenesis ideal for studies such as enhancer and gene traps, but doesn't allow targeting of specific sites in the genome. The ability to generate zebrafish transgenics for target genes has largely been restricted to inserting predicted regulatory elements, typically through bacterial artificial chromosome transgenesis (Suster et al., 2009).

For many years, gene targeting by homologous recombination has been used very effectively in mouse embryonic stem cells to specifically modify the mouse genome (Capecchi, 2005). This method relies on providing cells with exogenous DNA containing the insertion cassette flanked by arms containing a high degree of homology to the target region in the genome. Whilst this process is highly inefficient, inserting a selectable maker and working with a large number of totipotent cells have provided mouse geneticists with the tools to manipulate the genome (Kim and Kim, 2014). Unfortunately, until recently, the inefficiency of homologous recombination means that this groundbreaking technique is not suitable for organisms without an established embryonic stem cell line, such as zebrafish.

Armed with the knowledge that inducing double stranded breaks increases the rate of homologous recombination (Rouet et al., 1994), the advent of targetable nuclease technology is bringing us into an era where genome engineering is becoming applicable for a range of different organisms. In the previous chapter I described the use of TALENs to create mutations with a high efficiency in the

zebrafish genome. To date there have been two successful reports of homologous recombination using TALENs in zebrafish, one using short ssDNA as a donor (Bedell et al., 2012) and the other using larger dsDNA DNA (Zu et al., 2013). Although both of these methods represent a important leap in genetic modification in zebrafish, the relatively low efficiency of germline transmission (1.5% reported by Zu et al., 2013) remains a significant drawback.

An alternative method of inserting custom sequences at a targeted genomic location in cultured cells has been described that relies on the precise ligation of the overhangs produced by ZFNs or TALENs (Orlando et al., 2010, Miller et al., 2011, Maresca et al., 2013). This method, known previously as Obligate Ligation-Gated Recombination (ObLiGaRe) works through the dominant NHEJ pathway and allows the introduction of large cassettes with a much higher efficiency that homologous recombination *in vitro* (Maresca et al., 2013). I sought to adapt this strategy in order to insert custom cassettes into the zebrafish genome.

In this chapter I will describe the strategy for inserting exogenous DNA into the genome with the previously described TALENs. Using this technique, I show that inserting eGFP at the *plzfa* locus in zebrafish is efficient and precise, but results in the presence of indels around the target site and is highly mosaic in individual embryos. I implement a method to select embryos that contain the highest degree of insertion in order to improve germline transmission and discuss attempts made to improve the precision of insertion and prevent the donor vector backbone being inserted into the genome. Finally, I use this technique to further explore the expression and function of Plzf in the zebrafish hindbrain.

6.2 Strategy for inserting DNA sequences into a targeted locus

The previously described method (Maresca et al., 2013) makes use of a plasmid (hereafter called the 'donor plasmid'), which contains the TALEN or ZFN recognition site upstream of the insertion cassette. I elected to make use of the *plzfa* TALENs because of the evidence that they induce double stranded breaks at an efficiency approaching 100%, the gene's widespread expression pattern and the potential for further understanding of its biological function. My initial aim was to

insert a sequence encoding the eGFP protein into the *plzfa* locus, putting eGFP expression under control of the endogenous *plzfa* transcriptional machinery.

Donor plasmid design

Following the previously published guidelines, I designed a donor plasmid that contained the inverted *plzfa* left and right TALEN binding sites upstream of a promoter-less cassette containing P2A-eGFP followed by a polyadenylation signal. The self-cleaving P2A peptide is present to prevent Plzfa-eGFP fusion proteins from being generated that could have unknown biological effects. The reason for using inverted TALEN binding sites is to prevent the heterodimeric TALEN pair from being able to be recut the integrated DNA (Maresca et al., 2013). To allow future recombinase-mediated cassette exchange, the phiC31 integrase attP attachment sites flanked this cassette (Hu et al., 2011). Details of the generation of the *plzfa* donor plasmid can be found in Chapter 2.8.

I designed the donor plasmid with an emphasis on simplicity for future use. Complementary single stranded oligonucleotides containing the inverted TALEN binding site for a gene of choice can be annealed and cloned into the plasmid to replace the *plzfa* TALEN sites in a single step. Furthermore, restriction enzyme sites flank the eGFP cassette such that simple cloning strategies can be used to insert alternative sequences into the target locus. The annotated sequence for the *plzfa* donor plasmid is shown in Appendix Figure 8-6**.**

Insertion strategy

The theory behind the insertion technique is that upon injection of both the *plzfa* TALEN RNA and the *plzfa* donor plasmid DNA into the zebrafish embryo, the TALENs will induce double stranded breaks in both the genome and the donor plasmid (Figure 6-1a & b). Previous evidence has shown that these breaks produce complementary overhangs and will lead to the donor plasmid being precisely ligated into the *plzfa* locus of the zebrafish genome (Maresca et al., 2013). This will therefore put eGFP expression under the control of the endogenous *plzfa* transcriptional machinery, labelling cells that contain the insert with the fluorescent protein (Figure 6-1c). The presence of the P2A peptide means that eGFP will only be translated if the insertion event happens in frame with the Plzfa protein.

Figure 6-1: Strategy for DNA insertion

Cartoon depicting *in vivo* ligation event.

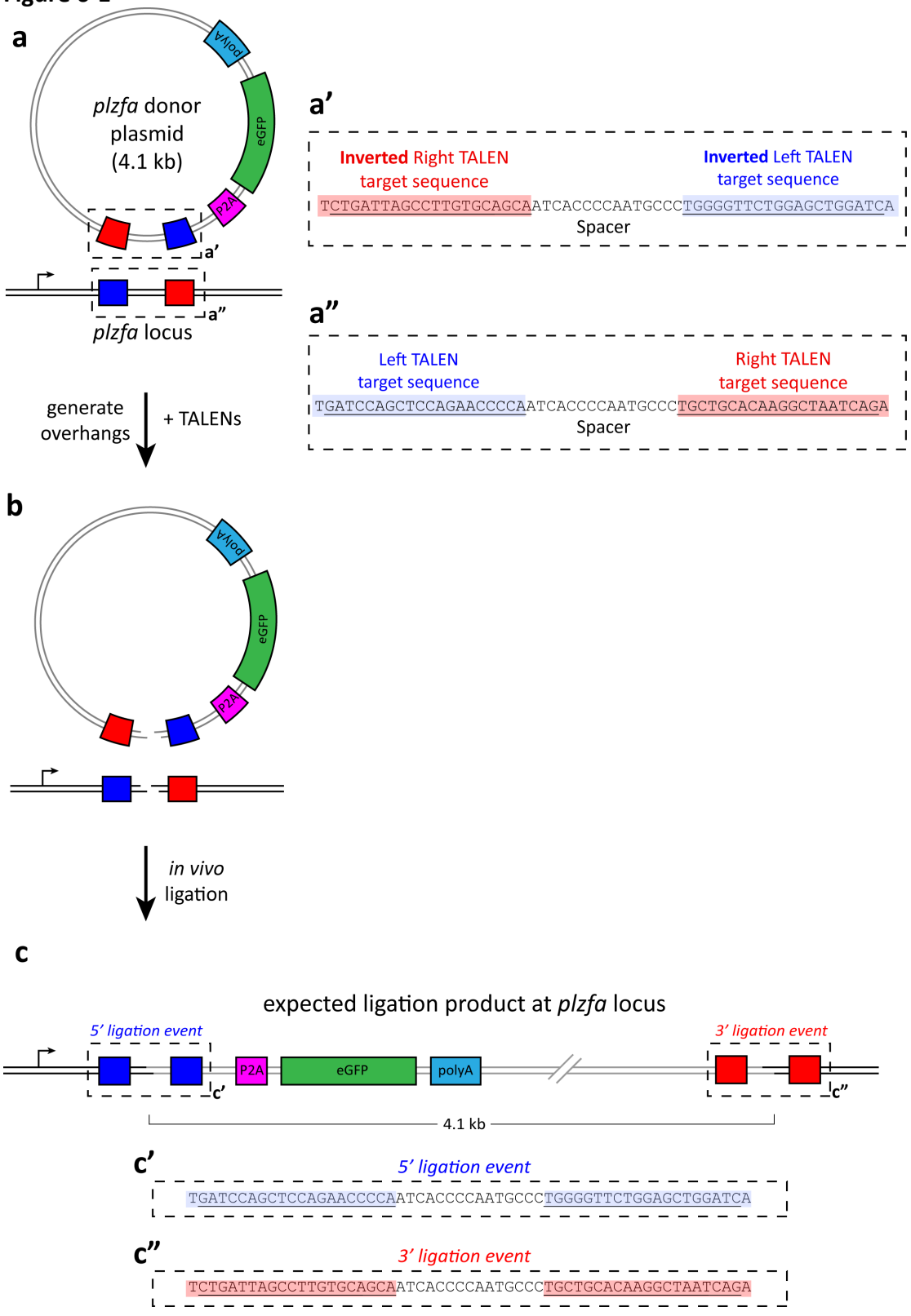
a: The *plzfa* donor plasmid contains the inverted left and right *plzfa* target sequences upstream of the P2A-eGFP-polyA insertion cassette. The forward strand sequences for TALEN target sequences in the plasmid (a') and genome (a") are shown. Note that although the target sequences are inverted, the spacer remains in the same orientation. Sequence for the donor plasmid is shown in Appendix Figure 8-6.

b: The addition of TALENs results in complementary overhangs being formed in the plasmid and the *plzfa* locus in the genome.

c: The overhangs anneal, inserting the entire 4.1 kb *plzfa* donor plasmid into the *plzfa* locus. The forward strands for the expected *5'* and *3'* ligations events are shown (c' and c" respectively).

Not drawn to scale.

Figure 6-1


a
plzfa donor plasmid (4.1 kb)
polyA
eGFP
P2A
a'
a"
plzfa locus
generate overhangs
+ TALENs
a'
Inverted Right TALEN target sequence
Inverted Left TALEN target sequence
TCTGATTAGCCTTGTGCAGCAATCACCCCAATGCCCTGGGGTTCTGGAGCTGGATCA
Spacer
a"
Left TALEN target sequence
Right TALEN target sequence
TGATCCAGCTCCAGAACCCCAATCACCCCAATGCCCTGCTGCACAAGGCTAATCAGA
Spacer
b
in vivo ligation
c
expected ligation product at plzfa locus
5' ligation event
3' ligation event
P2A
eGFP
polyA
c'
c"
4.1 kb
c'
5' ligation event
TGATCCAGCTCCAGAACCCCAATCACCCCAATGCCCTGGGGTTCTGGAGCTGGATCA
c"
3' ligation event
TCTGATTAGCCTTGTGCAGCAATCACCCCAATGCCCTGCTGCACAAGGCTAATCAGA

The ligations events formed in the expected product are resistant to further cutting by the heterodimeric TALENs as described previously (Figure 5-1b).

6.3 GFP is integrated into the *plzfa* locus with a high efficiency

I observed eGFP expression in over 75% of embryos injected with *plzfa* TALENs and the donor plasmid together, whilst injection of donor plasmid alone did not lead to any specific eGFP signal (Figure 6-2a-d). Encouragingly, the eGFP positive cells were found where the *plzfa* transcript is normally expressed, such as within the developing nervous system (Figure 6-2e-f). In order to determine whether the insertion event is specific to the *plzfa* locus, I injected the TALENs and donor plasmid into the previously described *plzfa* TALEN F1 embryos, which contain indels around the *plzfa* locus (Figure 5-8). I observed no specific eGFP expression in these embryos (Figure 6-2k-m). I found that increasing the amount of DNA that was injected resulted in an increase in embryo deformity and death whilst not affecting integration efficiency (Figure 6-2n & o) and therefore injected 27 pg or less DNA for all future experiments.

I designed primers to amplify the *5'* and *3'* insertion events at the *plzfa* locus by PCR of the genomic DNA from individual injected embryos and observed a band indicating successful integration in those embryos with eGFP expression (Figure 6-2y). Sequencing this insertion event revealed that whist the cassette was integrated into the genome, this was always accompanied by small insertions or deletions around the spacer region (Figure 6-2z), similar to what is seen during NHEJ-mediated DNA repair. To test whether this imprecise insertion was specific to the targeted *plzfa* locus, we used the same approach to insert eGFP at the *golden* locus using the previously described TALENs (Appendix Figure 8-7). I confirmed that eGFP is successfully integrated into the *golden* locus and found that sequencing revealed a similar extent of imprecision around the ligation event (Appendix Figure 8-7).

Figure 6-2: Evidence for efficient integration of GFP into the plzfa locus

a-d: Embryos injected with *plzfa* donor plasmid alone (a & b) had no GFP expression (23/23) whilst those injected with *plzfa* TALENs along with the plasmid showed GFP expression (92/119). Scale bar = 100 µm.

e-m: Injected embryos were analysed by confocal microscopy at 24 hpf after IHC with an anti-GFP antibody. GFP expressing cells were observed within the hindbrain and otic vesicle, where *plzfa* transcripts are normally detected in wild type embryos injected with both TALENs and donor plasmid (h-j). *plzfa* -/- embryos were injected with TALENs and donor and showed no GFP expression (k-m, 30/30). Scale bar = 20 µm.

n & o: Increasing amounts of donor plasmid were injected into the embryos and both their GFP expression (k) and survival (l) were scored at 24 hpf. Increasing the amount of DNA injected did not improve the efficiency of GFP expression but did decrease the survival of the embryos.

Figure 6-2

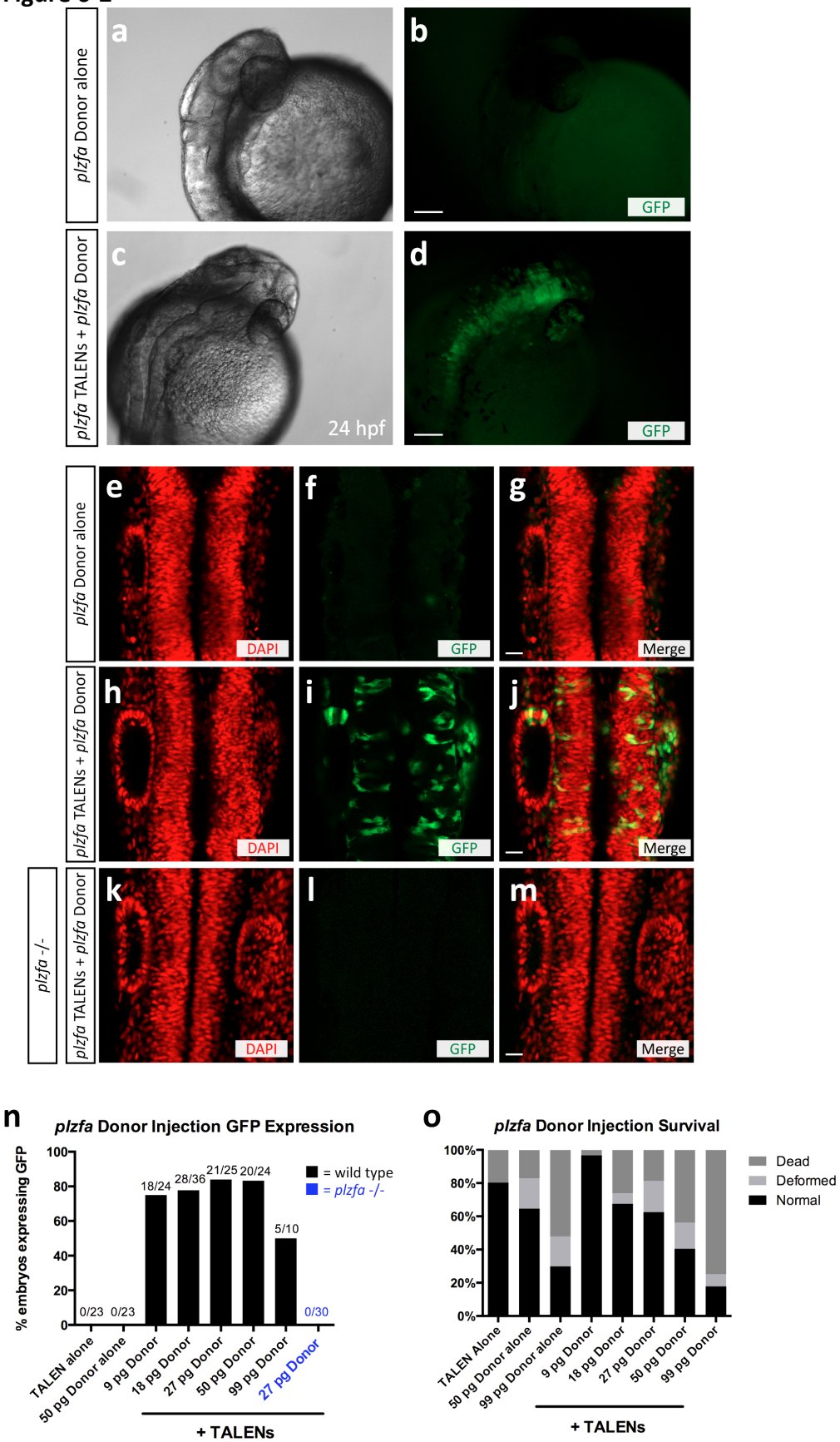

plzfa Donor alone
plzfa TALENs + plzfa Donor
plzfa -/-
a
b
c
d
e
f
g
h
i
j
k
l
m
24 hpf
GFP
DAPI
Merge
n
plzfa Donor Injection GFP Expression
% embryos expressing GFP
= wild type
= plzfa -/-
0/23
0/23
18/24
28/36
21/25
20/24
5/10
0/30
TALEN alone
50 pg Donor alone
9 pg Donor
18 pg Donor
27 pg Donor
50 pg Donor
99 pg Donor
27 pg Donor
+ TALENs
o
plzfa Donor Injection Survival
Dead
Deformed
Normal
TALEN Alone
50 pg Donor alone
99 pg Donor alone
9 pg Donor
18 pg Donor
27 pg Donor
50 pg Donor
99 pg Donor
+ TALENs

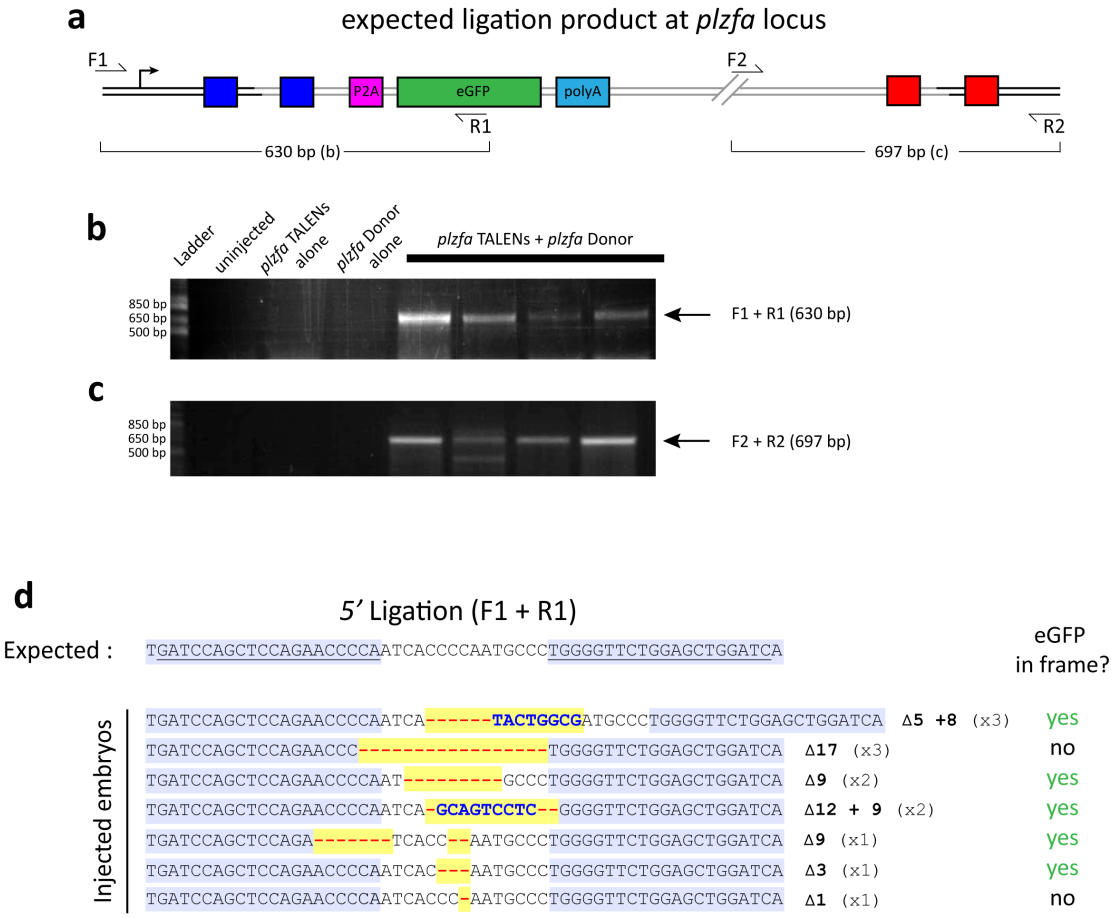
Figure 6-3: Insertion is associated with indels

a: Expected ligation product at the *plzfa* locus. Positions of primers (F1, R1, F2 & R2) are shown, along with the sizes of the two PCR products (650 and 697 bp respectively). F1 and R2 bind sequences found in the genomic DNA whilst R2 and F2 bind the donor plasmid DNA.

b & c: A PCR band corresponding to the *5'* ligation (b) and the *3'* ligation (c) is observed in embryos injected with *plzfa* TALENs + *plzfa* Donor plasmid.

d: The *5'* ligation PCR product was sequenced. In individual embryos several different alleles were sequenced, showing a variety of insertions and deletions around the ligation site. None of the observed PCR products showed precise ligation, however some contained indels that kept the eGFP protein in-frame.

Figure 6-3



d

5′ Ligation (F1 + R1)

			eGFP in frame?
Expected :	TGATCCAGCTCCAGAACCCCAATCACCCCAATGCCCTGGGGTTCTGGAGCTGGATCA		
Injected embryos	TGATCCAGCTCCAGAACCCCAATCA------**TACTGGCG**ATGCCCTGGGGTTCTGGAGCTGGATCA	**Δ5 +8** (x3)	yes
	TGATCCAGCTCCAGAACCC----------TGGGGTTCTGGAGCTGGATCA	**Δ17** (x3)	no
	TGATCCAGCTCCAGAACCCCAAT---------GCCCTGGGGTTCTGGAGCTGGATCA	**Δ9** (x2)	yes
	TGATCCAGCTCCAGAACCCCAATCA-**GCAGTCCTC**--GGGGTTCTGGAGCTGGATCA	**Δ12 + 9** (x2)	yes
	TGATCCAGCTCCAGA-------TCACC--AATGCCCTGGGGTTCTGGAGCTGGATCA	**Δ9** (x1)	yes
	TGATCCAGCTCCAGAACCCCAATCAC---AATGCCCTGGGGTTCTGGAGCTGGATCA	**Δ3** (x1)	yes
	TGATCCAGCTCCAGAACCCCAATCACCC-AATGCCCTGGGGTTCTGGAGCTGGATCA	**Δ1** (x1)	no

6.4 Selecting embryos for insertion significantly improves germline transmission

In order to determine whether the insertion can be passed on through the germ-line, I allowed the injected embryos to develop to adulthood and began analysing their progeny by GFP expression and by PCR. None of the adult fish analysed were found to transmit the insertion (0/23, *data not shown*). Examining the embryos injected with both *plzfa* TALENs and the *plzfa* donor plasmid revealed that in-frame integration is highly mosaic; some embryos have widespread GFP expression whilst in others only a few cells can be observed (Figure 6-4a-d). This suggested that germ-line efficiency could be significantly improved by only growing up embryos that have been selected for the highest amount of integration.

In the current strategy, GFP is being expressed from the endogenous *plzfa* transcriptional machinery, resulting in very weak florescence that is difficult to detect by wide-field fluorescence. To make it simpler to select embryos, I replaced the eGFP cassette with one containing the transcriptional activator GAL4 and an amplifying reporter cassette (*plzfa* GAL4 donor). This amplifying reporter is composed of 5 copies of the GAL4 binding sites (*UAS*), followed by the E1b basal promoter and the fluorescent tdTomato protein (Figure 6-4e). Similar amplification cassettes have been used in the generation of enhancer trap lines and have been found to produce strong fluorescent signal (Distel et al., 2009). I found that injection of this plasmid along with the *plzfa* TALENs resulted in strong tdTomato expression allowing simple selection of embryos containing the highest degree of in-frame integration (Figure 6-4f-i).

A commonly used method to identify carriers during Tol2 transgenesis is to insert an internal reporter that contains the alpha crystalline promoter driving expression of a fluorescent protein within the lens of the embryos (Gerety and Wilkinson, 2011). Fluorescence can be observed from around 2 days of development and is clearly observed by 3 days. *Tol2*-mediated DNA integration in the lens has previously been associated with high somatic transposition throughout the embryo (Balciunas et al., 2006) and unpublished experiments from the Wilkinson lab have shown that selecting injected embryos with fluorescent eyes positively

correlates with high germline transmission. I therefore included this internal reporter within the eGFP and GAL4 donor plasmids (Figure 6-4e; eGFP ACR & GAL4 ACC) and selected injected embryos that contained red or green eyes (Figure 6-4j-m). This internal reporter is dependent upon its own promoter and therefore will be expressed in embryos where the insertion has happened, irrespective of whether it is in-frame or not.

Embryos injected with the *plzfa* eGFP ACR donor plasmid that were selected for the presence of red eyes were grown to adulthood and their progeny screened. I found that 2 out of 8 of the founders transmitted the insertion to their progeny, revealed by the presence of RFP fluorescence in the lens of the F1 embryos (Figure 6-5a). No eGFP was observed in these embryos, suggesting that the insertion was out-of-frame, and this was confirmed by PCR and sequencing (Figure 6-5f). Embryos injected with *plzfa* GAL4 ACC donor plasmid and selected for high tdTomato expression were found to also transmit the insertion to the progeny. Out of 19 tested, 5 transmitted the insertion and 1 of these transmitted insertions that are in-frame. This was determined by observing tdTomato expression throughout the CNS of embryos produced from this founder, as well as citrine expression in the lens (Figure 6-5b-e), and was confirmed by sequencing (Figure 6-5f).

In order to confirm that the transgenic lines produced accurately reflect the *plzfa* expression patterns, I crossed the *Tg(plzfa:GAL4; UAS:tdTomato; ACC)* founder fish with a previous generated transgenic line *Tg(UAS:H2B-Citrine)* (Nikolaou et al., 2009), which will drive nuclear citrine expression in the GAL4 expressing cells. I examined expression at 12 hpf, where the *plzfa* transcript expression is absent from the anterior hindbrain (Figure 4-2a & Figure 6-5g). I observed H2B-Citrine expression in the same regions as Plzfa protein is expressed, indicating that the transgenic line does accurately recapitulate *plzfa* expression (Figure 6-5g-i). Comparing expression in the hindbrain at 24 hpf shows both Plzfa and H2B-Citrine expression throughout the hindbrain and developing otic vesicle (Figure 6-5j-l). In clutches of embryos obtained from individual founders, the number of embryos transmitting the transgene ranged from 2% (1/43 embryos) to 38% (21/56 embryos). The germline data is summarised in Figure 6-5m.

Figure 6-4: Methods of positive selection

a-d: High degree of mosaicism of insertion in embryos. 4 representative embryos showing GFP expression in the 24 hpf hindbrain. Scale bar = 20 µm.

e: Cartoon showing changes made to the *plzfa* donor plasmid to aid with embryo selection. The eGFP coding sequence was exchanged for a GAL4 cassette containing GAL4 – polyA, followed by 5 x *UAS* and tdTomato under the control of the E1b promoter. This plasmid becomes '*plzfa* GAL4 Donor'. The internal reporter alpha crystalline RFP or alpha crystalline citrine (ACR & ACC respectively) was inserted into the donor plasmids. These became *plzfa* ACR Donor and *plzfa* GAL4 ACC Donor respectively. Sequences for these cassettes are shown in Appendix Figure 8-8.

f-i: Example of living embryos at 72 hpf that were injected with *plzfa* TALENs and *plzfa* GAL4 Donor, showing widespread tdTomato expression (g & i).

j-m: Example of embryos injected either with *plzfa* TALENs and *plzfa* ACR Donor showing red eyes (k), or injected with *plzfa* TALENs and *plzfa* GAL4 ACC Donor showing green eyes (m).

Figure 6-4

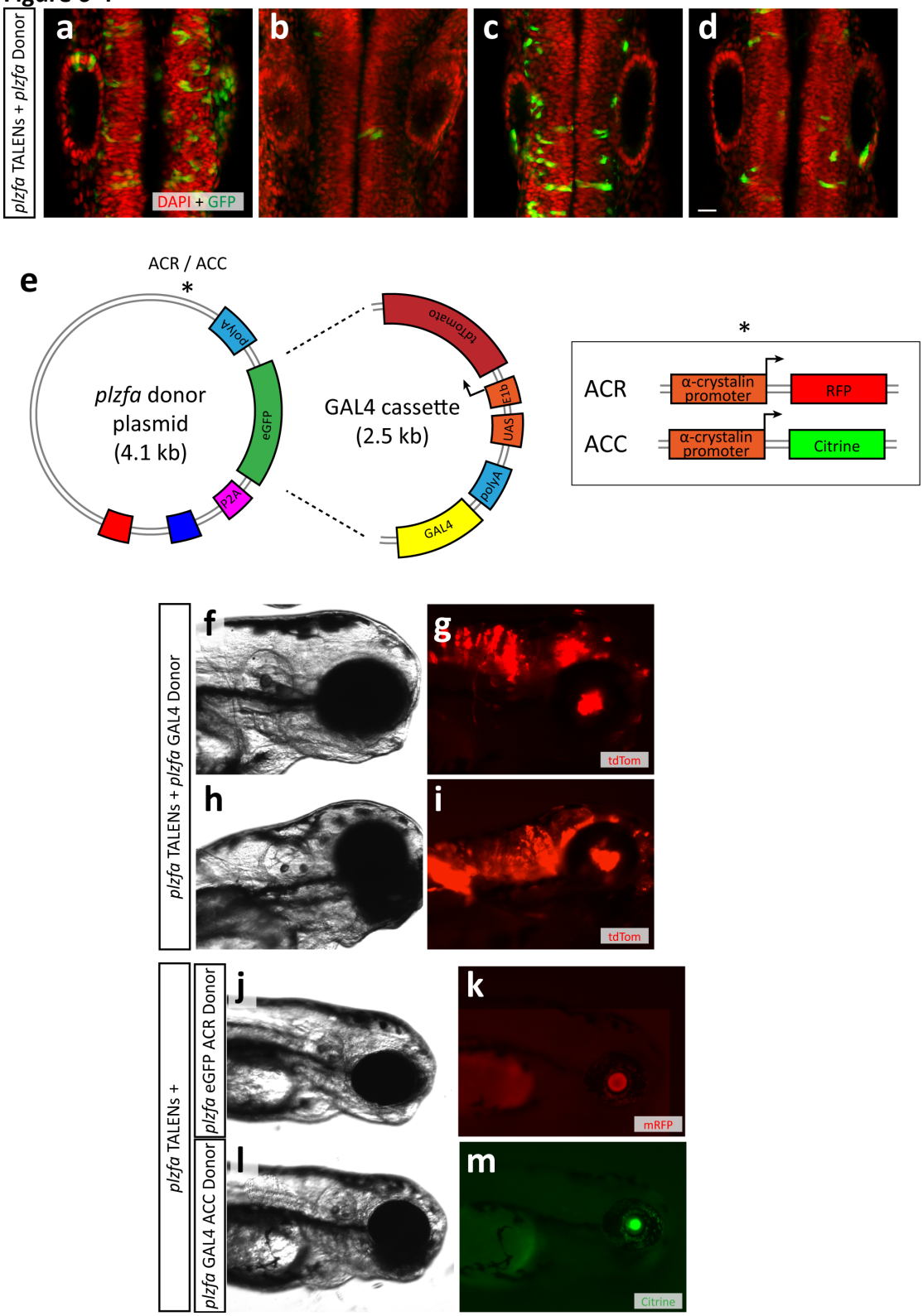

plzfa TALENs + plzfa Donor
a
DAPI + GFP
b
c
d
e
ACR / ACC
*
plzfa donor plasmid (4.1 kb)
polyA
eGFP
p2A
tdTomato
E1b
UAS
polyA
GAL4
GAL4 cassette (2.5 kb)
*
ACR
α-crystalin promoter
RFP
ACC
α-crystalin promoter
Citrine
plzfa TALENs + plzfa GAL4 Donor
f
g
tdTom
h
i
tdTom
plzfa TALENs +
plzfa eGFP ACR Donor
j
k
mRFP
plzfa GAL4 ACC Donor
l
m
Citrine

Figure 6-5: Germline transmission of insertion

a: Example of 3 dpf F1 embryos obtained from a *plzfa* ACR injected founder fish crossed to wild type. These embryos didn't display any eGFP expression (*data not shown)*, but several did contain mRFP expression in the lens. This indicates that insertion has occurred, but isn't in-frame.

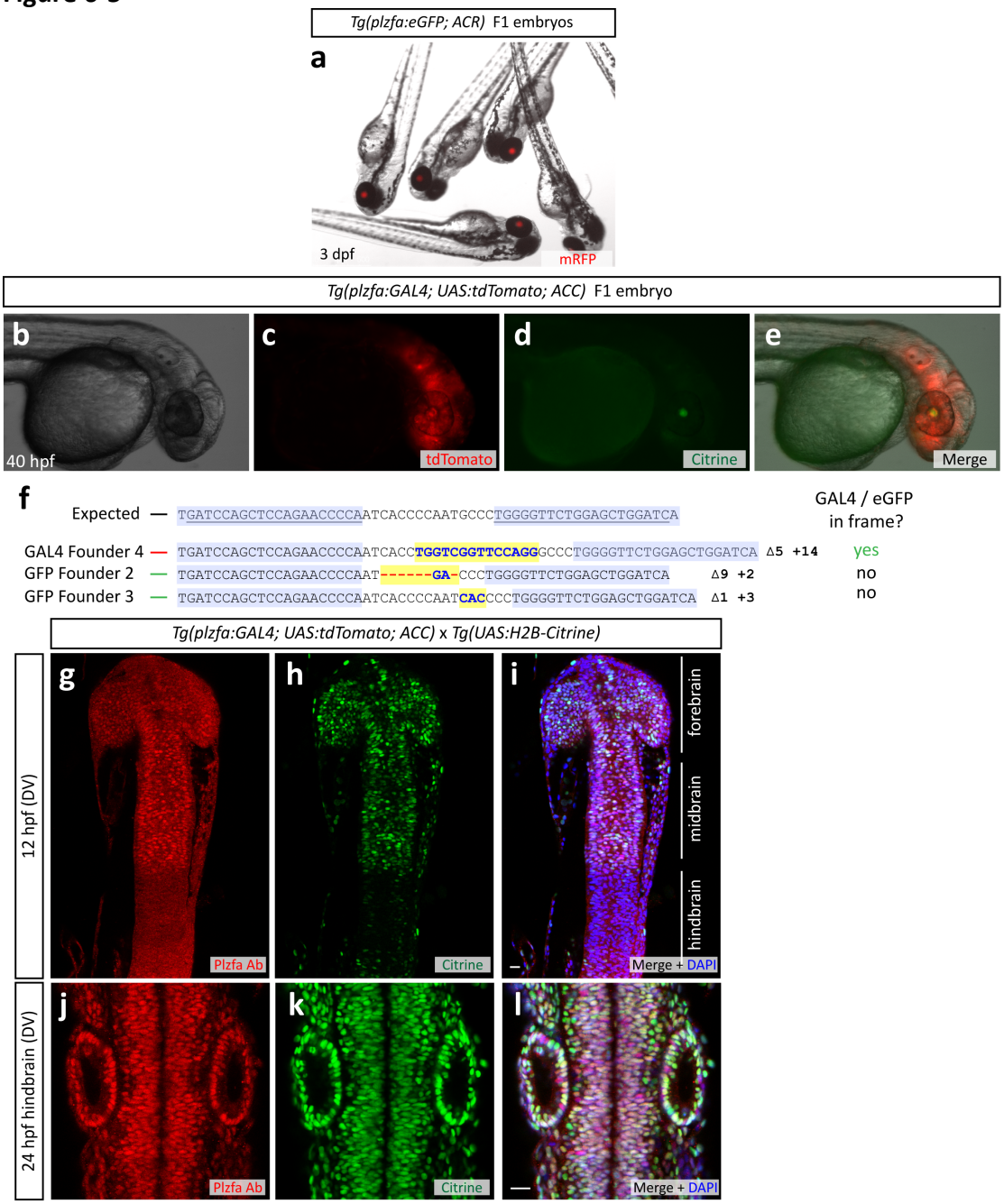
b-e: Example of a 40 hpf F1 embryo obtained from a *plzfa* GAL4 ACC injected founder fish crossed to wild type. This embryo shows tdTomato expression through the CNS (c) and citrine expressed in the lens (d), indicating that the insertion is present and is in-frame.

f: Sequencing results obtained from individual F1 embryos obtained from the indicated F0 founder fish.

g-l: The *Tg(plzfa:GAL4; UAS;tdTomato; ACC)* founder fish identified was crossed to the transgenic line *Tg(UAS:H2B-Citrine)* and embryos analysed by IHC. At 12 hpf, Plzfa protein expression is absent from the anterior hindbrain (g). A similarly restricted expression pattern is observed for the H2B-Citrine expression (h). At 24 hpf, both Plzfa expression (j) and H2B-Citrine (k) are present through the hindbrain and otic vesicle. Scale bar = 20 μm.

m: Summary of data obtained from germline transmission.

Figure 6-5



m

Insertion Germline Transmission				
Donor Plasmid	Selection Method	Total founders screened	Total founders with insertion	Total founders with insertion in-frame
plzfa eGFP	None	23	0 (0%)	0
plzfa eGFP ACR	Red eyes	8	2 (25%)	0
plzfa GAL4 ACC	tdTomato	19	5 (26%)	1

6.5 Strategies for improving the technique

Improving insertion precision

I attempted several methods to improve the precision of the insertion event with the aim to ensure the insertion events occur in-frame. The previously published results *in vitro* suggested that the majority of ligation events were occurring precisely (Maresca et al., 2013) and therefore I examined the method further to determine whether there were any differences that could account for these results. One difference is the mutations in the *FokI* nuclease that ensures the TALENs are heterodimeric. The variant that I use contains a single mutation; R487D (DD) and D384R (RR) for the left and right TALENs respectively (L_DD + L_RR), whilst the published results use double mutants; E490K, I538K (KK) and Q486E, I499L (EL). I aimed to determine whether using these different *FokI* mutants would affect the accuracy of insertion.

I inserted the left and right *plzfa* TALEN binding sites into the KK and EL TALEN backbone respectively (L_KK + R_EL) and injected these into the zebrafish embryo. Performing HRM around the *plzfa* target site reveals a similar deviation from the uninjected melt curves as observed for the L_DD + R_RR injections (Figure 6-6a). I used these double mutant TALENs to insert GFP at the *plzfa* locus, observing GFP expression in the majority of embryos (52/75 injected). Sequencing these embryos revealed a similar extent of insertions and deletions around the integration site as observed with the DD/RR TALENs (Figure 6-6b). I also confirmed that a pair of arrays which both contained the DD mutation (L_DD + R_DD) was unable to induce mutations (Figure 6-6a), suggesting that the indels observed are not caused by the TALENs being able to bind the ligation product.

Previous experiments to limit the extent of indels produced by targetable nuclease have involved generating a 'nicking' version of the nuclease (nickase) that induces single stranded breaks rather than double stranded breaks (Wang et al., 2012, Ramirez et al., 2012, Kim et al., 2012, Ran et al., 2013). Paired nickases, which produce two single stranded breaks on different DNA strands, have been shown to precisely integrate donor DNA containing complementary overhangs (Ran et al.,

2013). I therefore aimed to determine whether I could use double nicking TALENs to more precisely insert exogenous DNA into the *plzfa* locus.

In order to generate a nickase, I introduced a mutation (D450A) known to render the *FokI* catalytically inactive (Sanders et al., 2009) in the RR TALEN backbone. When this inactive array (rr) heterodimerises with the active DD array, a single stranded break with be induced (Figure 6-6c). Inserting two pairs of nickases (L_DD + R_rr + R_DD + L_rr) will generate a double-stranded break as previously described (Kim et al., 2012). The extent of indel formation produced by injection of RNA encoding the 4 arrays together was analysed by HRM (Figure 6-6d). The melt curves were observed to slightly differ from wild type, suggesting that indels are produced but not as extensively as when using the wild-type TALENs (compare with Figure 6-6a).

I subsequently injected the double nicking TALENs together with the *plzfa* donor plasmid into the zebrafish embryo. I found that the efficiency of integration was lower than observed when using the wild type TALENs, observing integration in 1 out of 5 embryos tested. Sequencing the PCR product for the *5'* ligation event from this embryo revealed that the insertion is associated with insertions and deletions of nucleotides.

Removing the vector backbone

The described technique involves the insertion of the entire plasmid into the zebrafish genome. I aimed to determine whether I could insert the expression cassette into the correct locus without the vector backbone. In order to do this, I designed a second *plzfa* donor plasmid (*plzfa* donor plasmid v2) that contains two inverted *plzfa* TALEN binding sites flanking the cassette (Figure 6-7b). I hypothesised that the TALENs will cut at both sites in the plasmid as well as a single site in the genome, therefore inserting only the eGFP cassette into the *plzfa* locus (Figure 6-7b-e). In order to prevent the possibility of only a single site being cut in the *plzfa* donor plasmid v2 and insertion of the entire plasmid, I digested and purified a linear fragment of the plasmid (Figure 6-7c).

Figure 6-6: Changing the *FokI* nuclease

a: HRM curves for embryos injected with various combinations of *plzfa* TALENs. Both the DD + RR (red lines) and KK + EL (blue lines) TALEN injected embryos melt curves deviate from the uninjected curves (black lines) to a similar extent. The curves produced from embryos injected with Left and Right TALEN arrays both containing the DD *FokI* variant (L_DD + R_DD, orange lines) are indistinguishable from the uninjected curves.

b: Embryos injected with the *plzfa* donor plasmid and the L_KK + R_EL *plzfa* TALENs were sequenced around the *5'* insertion site. Indels are observed around the spacer region.

c: Cartoon demonstrating action of double nickases. L_DD + R_rr TALENs (i) are able to induce a single stranded break (nick) in the forward strand (ii). L_rr + R_DD (iii) TALENs bind and produce a nick in the reverse strand (iv). Addition of all 4 TALEN arrays together produces a double-stranded break.

d: HRM curves for embryos injected with both pairs of nicking TALENs. Injected embryos melt profiles (purple lines) slightly differ from the uninjected profiles (black lines), indicating the presence of some indels (7/13 embryos).

e: Embryos injected with the two pairs of nicking TALENs were sequenced around the *5'* insertion site. Indels are observed around the spacer region.

Figure 6-6

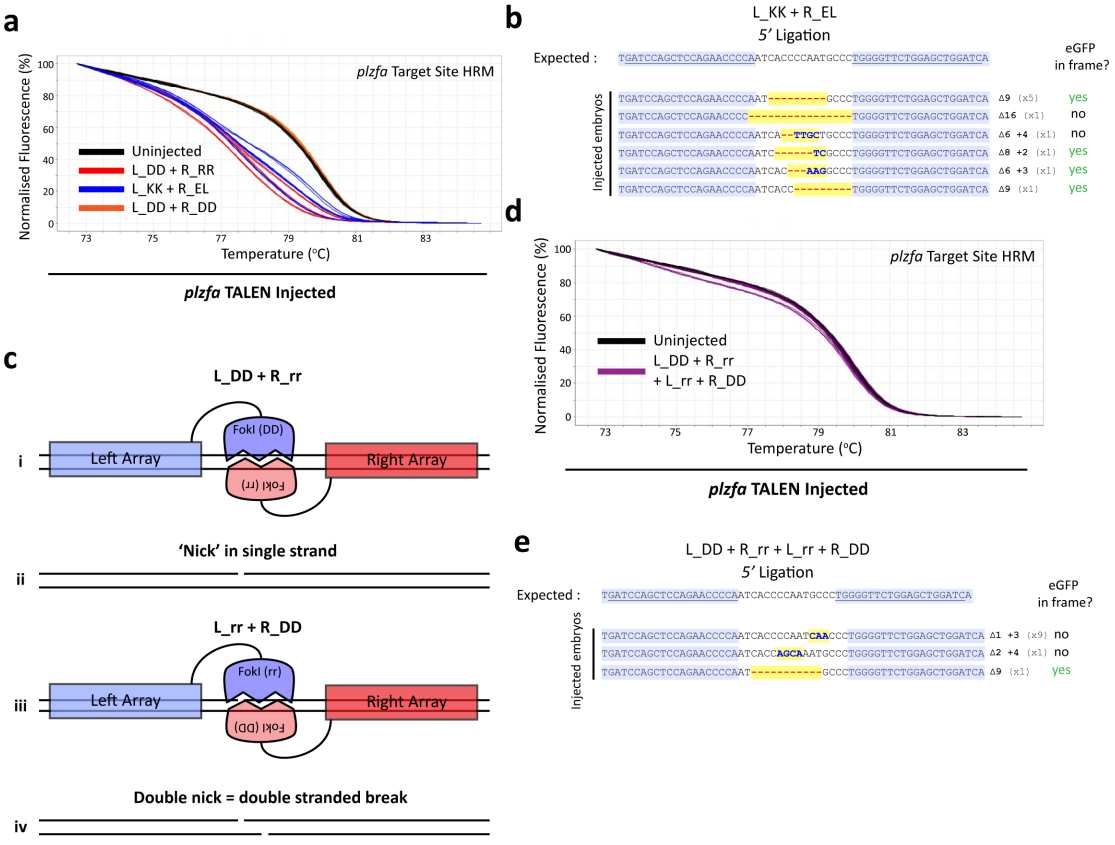

a
plzfa Target Site HRM
Normalised Fluorescence (%)
Temperature (°C)
Uninjected
L_DD + R_RR
L_KK + R_EL
L_DD + R_DD
plzfa TALEN Injected
b
L_KK + R_EL
5′ Ligation
Expected :
eGFP in frame?
Injected embryos
c
L_DD + R_rr
FokI (DD)
FokI (rr)
Left Array
Right Array
i
'Nick' in single strand
ii
L_rr + R_DD
iii
Double nick = double stranded break
iv
d
plzfa Target Site HRM
Uninjected
L_DD + R_rr
+ L_rr + R_DD
plzfa TALEN Injected
e
L_DD + R_rr + L_rr + R_DD
5′ Ligation

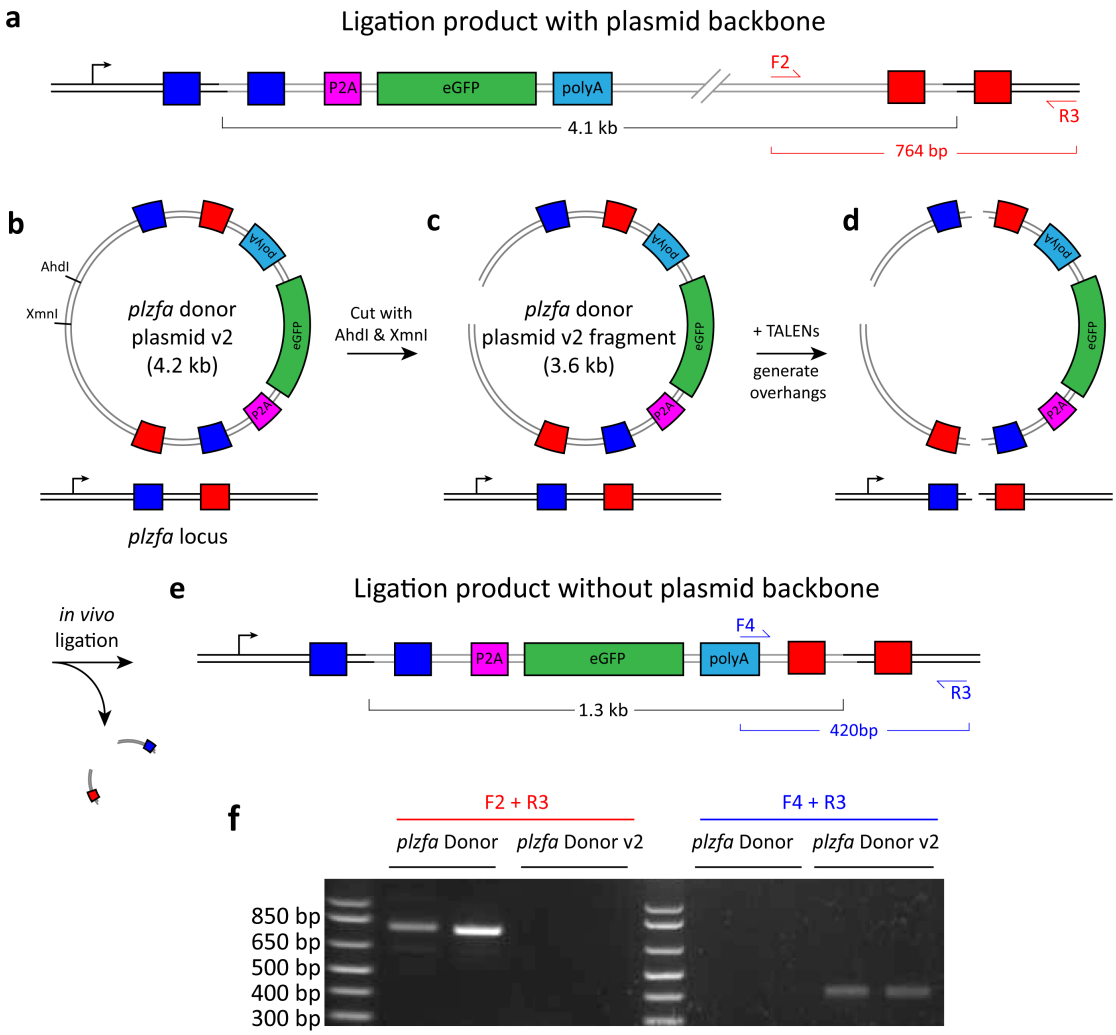
Figure 6-7: Insertion without the vector backbone

a: Ligation product at the *plzfa* locus with the vector backbone. Locations of primers to confirm insertion of vector backbone (F2 & R3) are shown in red.

b-e: Cartoon demonstrating theory behind insertions without the vector backbone. A second inverted *plzfa* TALEN binding site was inserted into the donor plasmid, such that the 2 pairs of sites flank the P2A-eGFP-polyA cassette (b). The plasmid was digested using the *AhdI* and *XmnI* restriction enzymes before injection and the resulting 3.6 kb linear fragment was purified (c). This fragment was injected along with the *plzfa* TALENs into the zebrafish, resulting in both sites being cut in the plasmid and a single site cut at the *plzfa* locus in the genome (d). The ligation event will insert the cassette into the *plzfa* locus without the vector backbone (e). Locations of the primers used to confirm this insertion (F4 & R3) are shown in blue. Not drawn to scale.

f: Gel showing insertion of the cassette without the vector backbone is possible. Two embryos injected with TALENs and the *plzfa* donor plasmid, or the *plzfa* donor plasmid v2, were taken for PCR analysis with primers that amplify ligation with or without the vector backbone (F2 + R3 and F4 + R3 respectively). F2 + R3 product was observed for *plzfa* Donor but not for *plzfa* Donor v2 injected embryos, whilst F4 + R3 product was observed for *plzfa* Donor v2 and not *plzfa* Donor injected embryos.

Figure 6-7


a
Ligation product with plasmid backbone
P2A
eGFP
polyA
F2
R3
4.1 kb
764 bp
b
AhdI
XmnI
plzfa donor
plasmid v2
(4.2 kb)
Cut with
AhdI & XmnI
c
plzfa donor
plasmid v2 fragment
(3.6 kb)
+ TALENs
generate
overhangs
d
plzfa locus
in vivo
ligation
e
Ligation product without plasmid backbone
F4
1.3 kb
420bp
f
F2 + R3
F4 + R3
plzfa Donor
plzfa Donor v2
plzfa Donor
plzfa Donor v2
850 bp
650 bp
500 bp
400 bp
300 bp

Individual embryos injected with TALENs and either *plzfa* donor plasmid or donor plasmid v2 were lysed and analysed by PCR for the presence of the backbone (Figure 6-7f). A PCR product for the vector backbone (F2 + R3) is observed for embryos injected with the *plzfa* Donor but not for most embryos injected with *plzfa* Donor v2 (4/6 no product observed). PCR product consistent with the integration of the cassette alone was observed in these embryos (F4 + R3 product). This preliminary data shows that the plasmid can be adapted for insertion without the vector backbone. Further experiments are required to determine whether this affects efficiency of the insertion event.

6.6 Using this technique to examine *plzfa* expression and function

Expression analysis

I carried out preliminary experiments to gain information about *plzfa* expression using the insertion technique. Embryos that were injected with *plzfa* TALENs and *plzfa* donor plasmid were grown to 44 hpf at which point they were processed for IHC. I compared GFP expression with differentiated neurons (HuC/D) and glial fibres (GFAP) and a representative embryo is shown in Figure 6-8. I consistently observed higher levels of GFP expression at locations flanking rhombomere boundaries (arrows in Figure 6-8a-c). As discussed previously, these cells are actively undergoing neurogenesis (Figure 4-6).

I observed some GFP expression in terminally differentiated neurons (arrowheads in Figure 6-8a'-c'). One possibility is that this GFP expression is reflective of *plzfa* expression in some neurons, although this was not observed in the ISH experiments (Figure 4-6). Alternatively, this signal could be caused by perdurance of the stable GFP protein that is still present despite *plzfa* transcription no longer occurring in these cells. Without further analysis, I am unable to determine which of these explanations is correct.

The previous data regarding expression of Plzf in the nervous system have shown it is expressed in neural progenitor cells and declines as cells become terminally differentiated (Elkabetz et al., 2008, Tailor et al., 2013, Sobieszczuk et al., 2010, Gaber et al., 2013). I have results that suggest both *plzfa* transcripts and Plzf protein are expressed in cells that are actively undergoing neurogenesis (Figure 4-5

& Figure 4-6). I therefore aimed to gain some insight into how *plzfa* transcription is being regulated in these cells.

I re-examined the expression of *plzfa* in the 44 hpf embryos by fluorescent ISH and observed higher levels of *plzfa* transcript in cells that are believed to have begun neurogenesis based upon their position flanking rhombomere boundaries (Figure 6-9a & b). A candidate gene for increasing *plzfa* expression in these differentiating cells is *neurog1* and I therefore aimed to explore whether *plzfa* is downstream of this proneural gene.

To do this, I injected embryos with *neurog1* MOs and analysed expression of *btbd6a* and *plzfa* (Figure 6-9c-f). I observed the previously described decrease in *btbd6a* expression in the morphants (Sobieszczuk et al., 2010), but did not observe any changes in *plzfa* expression. Secondly, I ectopically expressed *neurog1* by injecting RNA into the embryos along with H2B-Citrine as a marker and carried out ISH for *plzfa* (Figure 6-9g-h). The majority of injected embryos died during development, and the ones that survived contained mosaic H2B-citrine expression. In these embryos, a slight decrease in *plzfa* expression was observed in embryos containing H2B-Citrine expression. From these results I can't conclude whether *plzfa* expression is downstream of *neurog1*, as will be discussed further below.

Functional analysis

Observing both *plzfa* transcript and Plzf protein within the differentiating cells suggested that Plzfa could be having a role in these cells. Previous work showed that Plzfa acted to inhibit neurogenesis by inhibiting *neurog1*, and based upon the increased expression at the onset of differentiation, I was interested in determining whether Plzfa acts to inhibit genes involved in later steps of the neurogenic cascade. I aimed to test this by forcing cells to embark on the neurogenic cascade and determine whether ectopic expression of Plzfa could block later steps of differentiation.

I carried out preliminary experiments where I injected embryos with RNA *neurog1* RNA alone, or with RNA encoding Neurog1 and Plzfa, and carried out ISH for *neurod4* expression. I observe an increase in the number of *neurod4* expressing cells upon ectopic expression of *neurog1* (Figure 6-10d).

Figure 6-8: Analysis of *plzfa* expression levels in donor injected embryos at 44 hpf

Representative embryo at 44 hpf injected with *plzfa* TALENs and *plzfa* donor plasmid after IHC using anti-GFP (a), anti-GFAP (b) and anti-HuC/D (blue in c). Arrows indicate positions flanking rhombomere boundaries where neurogenesis is active. Higher magnification images are shown (a'-c'). Arrowheads mark cells that are double positive for HuC/D and GFP expression.

Scale bar = 20 μm

Figure 6-8

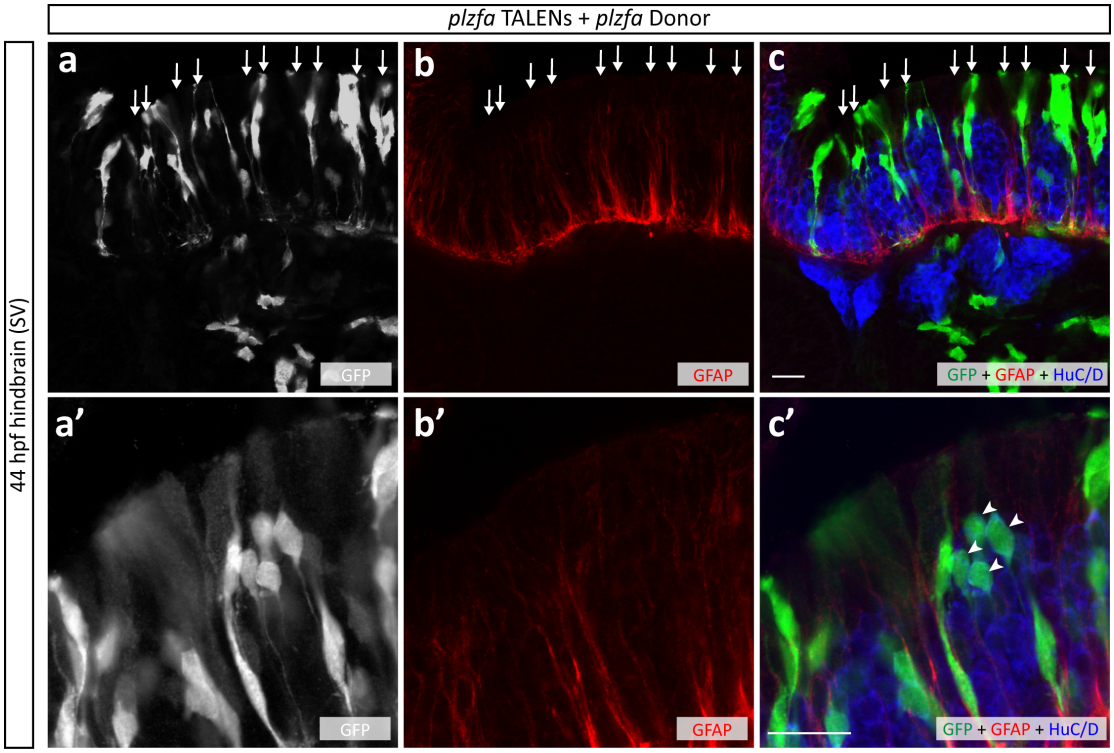

plzfa TALENs + *plzfa* Donor
44 hpf hindbrain (SV)
a
b
c
GFP
GFAP
GFP + GFAP + HuC/D
a'
b'
c'
GFP
GFAP
GFP + GFAP + HuC/D

Figure 6-9: Exploring whether *neurog1* regulates *plzfa* expression

a & b: Higher levels of *plzfa* transcript are observed at regions associated with neurogenesis. Fluorescent ISH of *plzfa* (a) is the same as displayed previously (Figure 4-6d') and was converted into a heatmap using ImageJ (b). Higher levels of expression are observed in cells flanking rhombomere boundaries (6/6 embryos).
c-f: Embryos were injected with either Control (c & e) or *neurog1* (d & f) MOs and analysed for *btbd6a* (c & d) or *plzfa* (e & f) expression by ISH at 12 hpf. Knockdown of Neurog1 leads to a major decreases in *btbd6a* expression (d, 15/15) but has no clear effect on *plzfa* expression (f, 0/19).
g-i: Uninjected embryos (g) were compared to those injected with 50 pg neurog1 and H2B-Citrine RNA (h & i) for *plzfa* expression. A slight decrease in *plzfa* expression was observed in cells containing ectopic RNA. This is most clearly observed at higher magnification (h') and when converted into a heatmap (h").
Scale bar = 20 μm

Figure 6-9

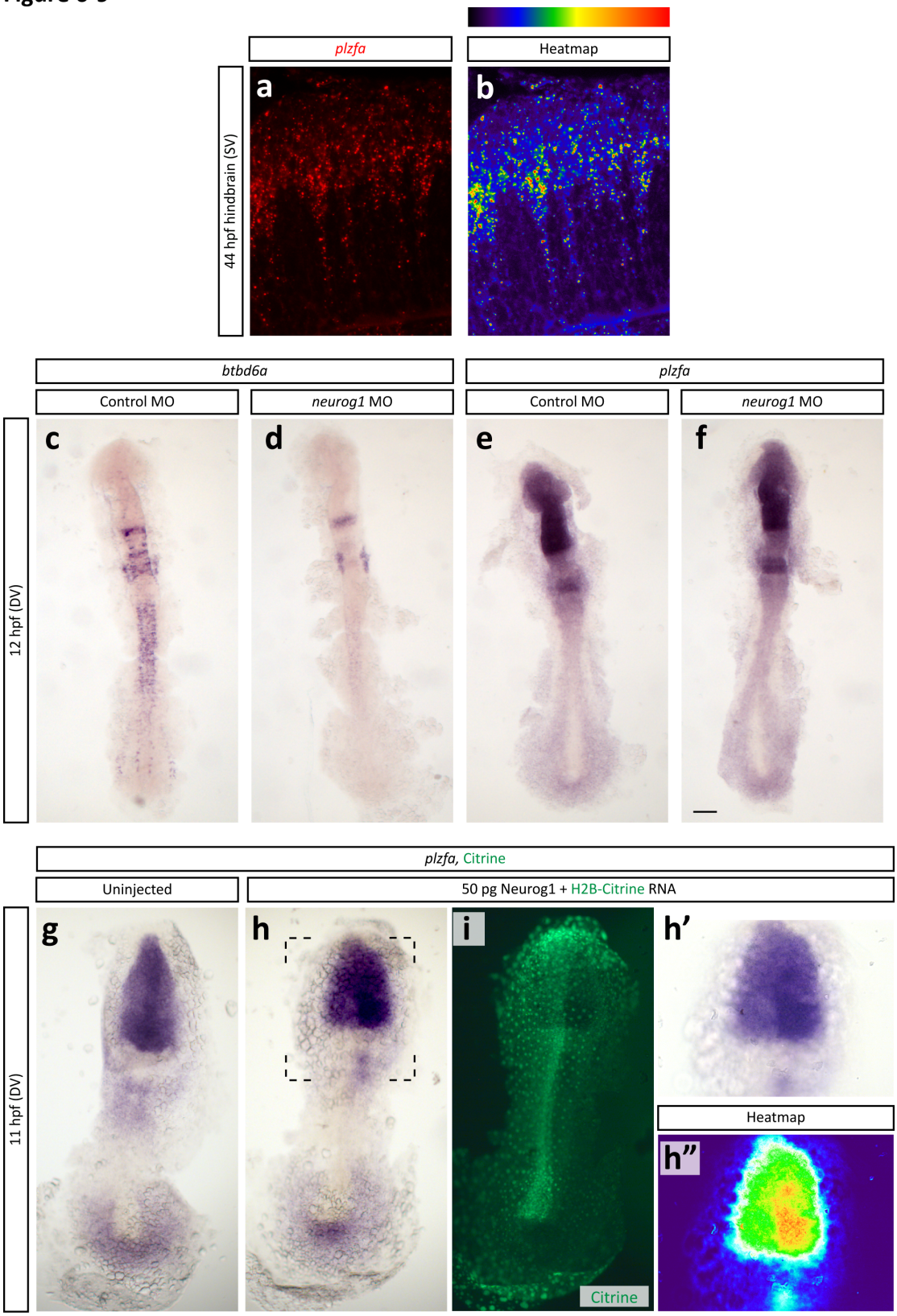
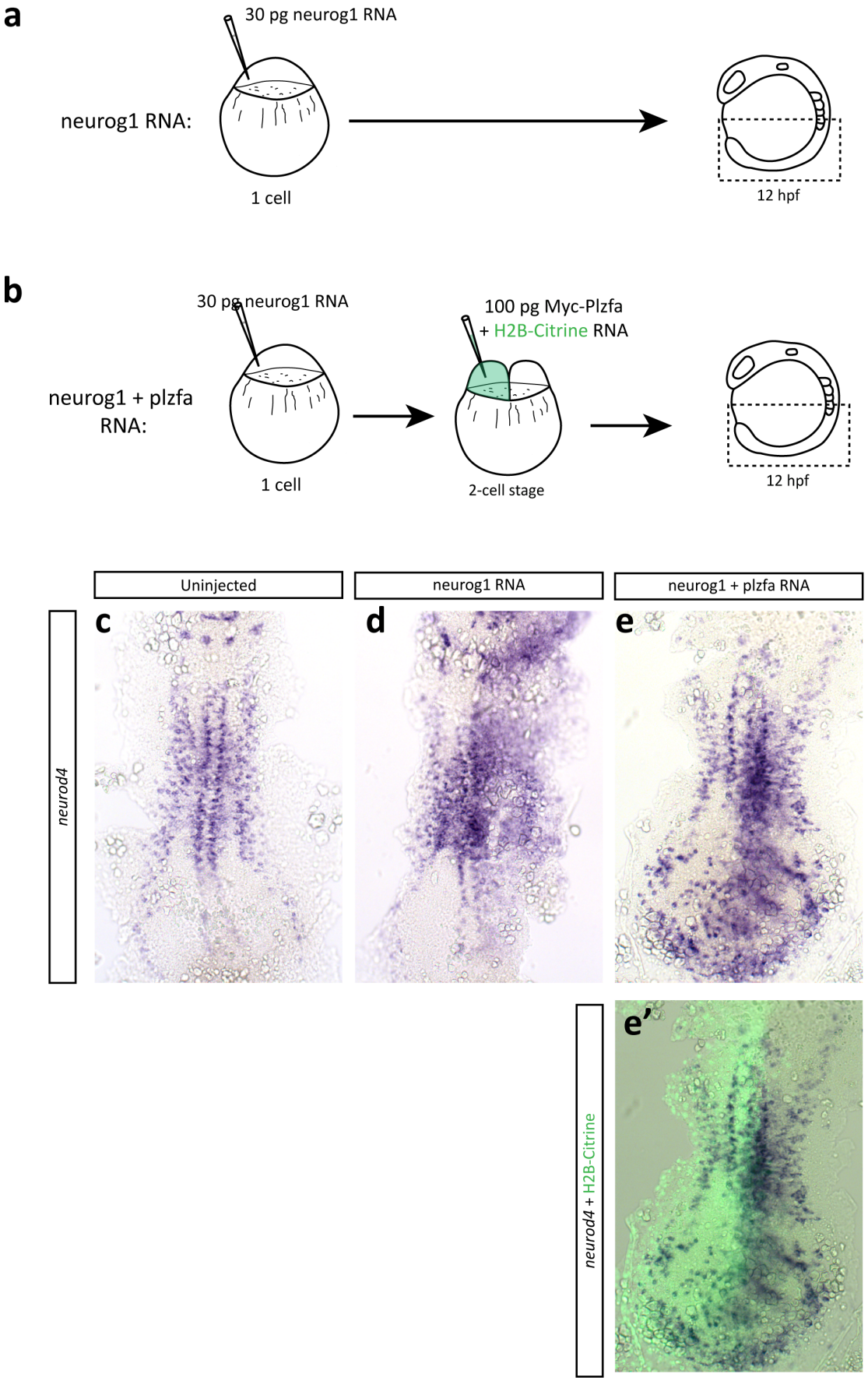

plzfa
Heatmap
44 hpf hindbrain (SV)
a
b
btbd6a
plzfa
Control MO
neurog1 MO
Control MO
neurog1 MO
12 hpf (DV)
c
d
e
f
plzfa, Citrine
Uninjected
50 pg Neurog1 + H2B-Citrine RNA
11 hpf (DV)
g
h
i
h'
Heatmap
h"
Citrine

Figure 6-10: Preliminary evidence that Plzfa inhibits multiple steps of neuronal differentiation

Embryos injected with 30 pg *neurog1* RNA at the one cell stage were allowed to develop until 12 hpf (a) and processed for *neurod4* ISH (d). A subset of these embryos were also injected with 100 pg *myc-plzfa* together with *H2B-citrine* RNA in one cell at the 2-cell stage (b). Theses embryos were similarly allowed to develop and processed for *neurod4* ISH at 12 hpf (e). Uninjected embryos were included as a negative control (c).

A total of 21 uninjected embryos were observed to have wild-type levels of *neurod4* (c). Embryos injected with *neurog1* RNA alone showed ectopic *neuod4* expression (d) in 12/16 embryos. Embryos injected with both *neurog1* (1-cell stage) and *myc-plzfa* (2-cell stage) resulted in reduced *neurod4* expression in 33/50, increased *neurod4* in 9/50 and wild-type levels of *neurod4* in 8/50 embryos.

Figure 6-10


a
30 pg neurog1 RNA
neurog1 RNA:
1 cell
12 hpf
b
30 pg neurog1 RNA
100 pg Myc-Plzfa
+ H2B-Citrine RNA
neurog1 + plzfa
RNA:
1 cell
2-cell stage
12 hpf
Uninjected
neurog1 RNA
neurog1 + plzfa RNA
neurod4
c
d
e
e'
neurod4 + H2B-Citrine

I subsequently took some embryos that had ectopic Neurog1 expression throughout the entire embryo and selectively overexpressed Plzfa in half of the embryo, by injecting *myc-plzfa* into one cell at the 2-cell stage. This allows Plzf function to be analysed in embryos known to have increased neurogenesis. Overexpressing *plzfa* appeared to block the increase in *neurod4* (Figure 6-10e). Although further experiments are required to confirm this result, this provides evidence that Plzf is acting at multiple steps in the neurogenic cascade to block differentiation.

6.7 Discussion

Insertions without homology arms

In this chapter I have described a novel method for the insertion of exogenous DNA into targeted locations in the zebrafish genome. I have demonstrated that insertions occur with a high efficiency in injected embryos and present evidence that the insertion is specific to the target locus. The procedure has been designed to be as simple as possible, whereby the target sequence can be replaced using a simple oligonucleotide cloning strategy and the insertion cassette replaced. Targeting two separate loci, I show that the ligation event is associated with the formation of indels resulting in frame-shifts and consequently the cassette won't be translated in all insertions.

Whilst this project was underway, another group published a similar technique for inserting DNA into the zebrafish using the CRISPR/Cas9 system (Auer et al., 2014). They showed that they were able to target *eGFP* that was stably integrated into the genome from previous *Tol2*-mediated transgenesis and replace it with other reporters. They report similar results regarding the formation of indels around the insertion site and show that positive selection of injected embryos resulted in a germline transmission rate up to 40%.

Interestingly, they show precise ligation of the insert in 17% of the injected embryos analysed (Auer et al., 2014), comparing to 0% in the embryos that I analysed. This may indicate that the CRISPR/Cas9 system can insert DNA with a greater precision than using TALENs. CRISPR/Cas9 is believed to leave blunt ends after cutting (Jinek et al., 2012, Garneau et al., 2010) whilst the *FokI* nuclease leaves 4 – 5 bp *5'* overhangs (Orlando et al., 2010). Further analysis would be needed to determine whether this difference in cutting accounts for changes in precision. Using a single CRISPR guide RNA to target DNA results in recreation of the target site in the resulting ligation product, an issue that is alleviated by using obligate heterodimeric TALENs as discussed (Maresca et al., 2013). It would therefore be expected that the CRISPR/Cas9 system would continue to generate indels in the ligation product, but the fact that they observe occasional perfect ligation suggests that this is not the case. They report that it is possible to insert DNA into an

endogenous locus (*kif5aa*), but they observe integration in only 9.6% of injected embryos and don't demonstrate germline transmission (Auer et al., 2014). This relatively low efficiency in injected embryos seems likely to be caused by the particular guide RNA being used to target *kif5aa*, which has an indel efficiency of 22%. This study demonstrates that this technique for insertion is applicable for all targetable nucleases available and suggests that insertion efficiency correlates with indel efficiency.

Importance of selection

My results suggest that selecting embryos that contain the highest degree of in-frame insertion results in the highest germline efficiency. I find that I cannot reliably screen embryos by inserting the coding sequence for the eGFP protein directly under control of the endogenous *plzfa* promoter, likely because the production of the transcript is too low. Instead, I elected to enhance the production of fluorescent proteins by inserting the GAL4 transcriptional activator and providing a fluorescent protein under control of the basal E1b promoter downstream of the *UAS* regulatory sequences. It has previously been shown that there is a delay in expression using this system, due to the time taken to produce the GAL4 protein followed by the fluorescent protein (Distel et al., 2009) and consequently the transgenic line may not recapitulate temporally accurate endogenous expression patterns. Secondly, there is known to be perdurance caused by the high half-life of the expressed proteins, such that fluorescence may still be observed in cells where the transcript is no longer expressed.

In this work, I show that by providing an internal reporter, such as a lens-specific promoter driving fluorescence, can be used as a selection method in injected embryos. Inserting this internal reporter eliminates the requirement for having to increase the fluorescence and therefore is suitable for experiments where non-fluorescent cassettes are inserted, or for targeting genes that don't have a widespread expression pattern. I have shown that after selection methods, germline transmission reaches efficiencies of around 25%. This is a significant improvement when compared to similar techniques that rely on homologous recombination, where rates between 1 and 4% have been reported (Bedell et al., 2012, Zu et al.,

2013). Furthermore, this method doesn't require detailed knowledge of the surrounding sequence of the target gene as is required for generating homology arms.

Improving the technique

A potential method to further improve transmission of the insert to the progeny is to target the integration event to the germ cells. Previous studies have shown this is possible by the addition of the *3'* UTR from the zebrafish *nanos1* gene, which enriches RNA bearing this sequence in the germline precursor cells (Köprunner et al., 2001). Use of this has previously been shown to significantly improve germline transmission for the recombinase mediated cassette exchange technique (Hu et al., 2011) and similar methods have been used in *Drosophila* to improve homologous recombination (Baena-Lopez et al., 2013). It would therefore be interesting to test whether targeting the TALEN pairs to the germ cells improves transmission.

I was unable to improve the precision of the insertion events by manipulating the *FokI* nuclease. It is interesting to note that the previous research in cultured cells found that the precision of the ligation event depended upon the cell type being used (Maresca et al., 2013). Cellular and genetic differences between the cell types may therefore influence ligation precision. In my experiments, the use of a self-cleaving P2A peptide necessitates that the insertion is in-frame with the targeted coding sequence in order to generate the protein being inserted. With the knowledge that the insertion is imprecise, the technique could be adapted such that in-frame insertion wouldn't be a requirement for translation. For example, the insertion could be targeted upstream of the start codon, or the donor plasmid could make use of an internal ribosome entry site (IRES) which has been used successfully in zebrafish genetics (Kwan et al., 2007). Flanking the insertion cassette with recombinase sites allows techniques such as recombinase-mediated cassette exchange to be carried out once I have a stable line (Hu et al., 2011). This technique could be used to both exchange the inserted transgenes with other cassettes, or could be used to correct situations where the gene has been inserted out-of-frame.

The FlipTrap technique in zebrafish has been developed to make use of fluorescent reporter constructs flanked with splice acceptor and splice donor sequences in order to generate fusion transcripts between the gene of interest and the reporter construct (Trinh et al., 2011). This technique was originally developed for Tol2-mediated integration, whereby the occasional random insertion into an intron results in a fusion between the reporter construct and surrounding exons. This has been used to generate fusion proteins with GFP, allowing real-time visualisation of endogenous proteins in the zebrafish. Combining FlipTrap with the technique described in this Chapter would allow researchers to generate fusions with the endogenous protein of their choice. The technique involves targeting gene introns and inserting splice acceptor and donor sites, and therefore any imprecision around the insertion site would not be detrimental to the in-frame translation of the fusion protein.

I present preliminary evidence to show that the donor plasmid can be adapted for insertion without the vector backbone. Whilst insertion of the backbone doesn't appear to be toxic for the zebrafish, there is evidence to suggest that this bacterial DNA can lead to gene silencing through histone modifications in mammalian cells (Riu et al., 2007). Previous studies have suggested that the addition of LoxP sites could be used to remove the backbone after stable integration (Maresca et al., 2013). I show that by implementing a second TALEN binding site in the donor plasmid, insertion of the cassette without the surrounding DNA is possible in injected embryos.

Investigations into plzfa expression

Analysis of *plzfa* transcript levels suggested that expression was increased in cells beginning to undergo neurogenesis. These cells are beginning to express *neurog1* and therefore I hypothesised that expression of *plzfa* would be induced downstream of this proneural gene. However, I found that sustained Neurog1 overexpression leads to a slight decrease in *plzfa* expression. In this experiment, I am forcing these cells to embark on neuronal differentiation and therefore this slight decrease may reflect the fact that these cells are at a point in the neurogenic cascade when *plzfa* expression is switched off. A more relevant experiment would

be to provide embryos with a brief burst of *neurog1* expression and determine whether this can induce *plzfa* expression. This would require more sophisticated techniques to temporally control induction of gene expression, such as heat shock or GAL4/UAS approaches (Esengil and Chen, 2008).

The finding that *plzfa* is expressed and may have a role in differentiating cells has not previously been reported for studies within the nervous system (Sobieszczuk et al., 2010, Gaber et al., 2013). The implications of this will be discussed further in the next chapter.

Conclusions

In this chapter I have established a novel technique for targeted insertions into the zebrafish genome using TALENs as an alternative to homologous recombination. I've shown that I am able to insert exogenous DNA into specific locations of the genome with a high efficiency and that these insertions are transmitted through the germline. I believe that this technique represents a significant addition to our growing repertoire of genome engineering tools in the zebrafish. Finally, although this work has focussed on the zebrafish, this technique is likely to be applicable for other model organisms.

7 Discussion

Plzf and progenitor maintenance

Generating the appropriate number and type of neural cells requires precise regulation of progenitor cell maintenance and cell differentiation. Plzf is widely expressed in neural progenitors but is largely restricted from post-mitotic differentiated neurons. This raises the possibility that it has an important role in progenitors. Previous data from the Wilkinson lab shows that Plzfa is therefore able to inhibit primary neurogenesis in the zebrafish embryo (Sobieszczuk et al., 2010). During normal development, Neurog1-induced upregulation of Btbd6a in cells undergoing neurogenesis removes Plzfa function. Consequently, removing Plzfa function through knockdown experiments has no effect upon the number of neurons produced during primary neurogenesis. Only if Notch activity is attenuated does the knockdown of endogenous Plzfa reveal a phenotype related to an inhibition of neurogenesis.

I explored the idea that the function of Plzfa was not entirely redundant to Notch-mediated lateral inhibition and asked whether redundancy between the two Plzf proteins in zebrafish was able to mask any phenotype related to the single knockdown in Chapter 4. This was based upon the finding that the mRNA of both paralogues was expressed in neural progenitors of the hindbrain. Knocking down both paralogues using MOs resulted in a phenotype that I attributed to a mild decrease in the number of neural progenitors present. This was evident at 44 hpf and revealed by a disruption to the pattern of differentiated neurons in the hindbrain. I was unable to determine whether this loss of progenitors was caused by premature differentiation or by decreased progenitor proliferation.

Due to concerns over the specificity of the MOs, together the recent development of TALEN technology, I aimed to explore this phenotype further by generating zebrafish embryos containing loss-of-function mutations in *plzfa* and *plzfb* in Chapter 5. Examining the same markers used for phenotypic analysis in the morphant embryos did not reveal any defect in progenitor maintenance. After exploring various possibilities, I propose that the phenotype I observed in the double morphants is a caused by Tp53-independent toxicity as a result of morpholino usage.

It is interesting to consider what could be causing the phenotype I observe in the MO knockdowns. I observe a global delay in development upon injection of the *plzfa* MO, irrespective of Tp53 function. I carried out careful controls to ensure that the embryos injected with either control or *plzfa* MOs were matched by developmental stage rather than time and therefore I don't believe that the delay can purely account for the observed phenotype.

The mechanism by which MOs are able to activate Tp53 activity is not known (Robu et al., 2007, Gerety and Wilkinson, 2011). The normal function of the Tp53 protein is to integrate multiple signalling pathways activated in response to stress signals and result in apoptosis of the affected cell (Murray-Zmijewski et al., 2006). It is therefore possible that the off-target affects associated with MO usage are caused by activation of stress pathways. Therefore, whilst blocking Tp53 function may prevent cell apoptosis, other outcomes of cell stress such as cell cycle arrest may be relevant in MO injected embryos and could account for the morphant phenotype. Preliminary experiments investigating this showed that thermal stress was unable to recreate the spurious phenotypes observed upon MO injection (Law and Sargent, 2014). Components of protein kinase cascades pathways downstream of stress stimuli should be analysed in the context of MO injection in order to investigate this concept further (Kyriakis and Avruch, 1996).

Based upon these results, it therefore seems likely that the original hypothesis that Plzfa function in preventing the onset of neurogenesis is entirely redundant to Notch-mediated lateral inhibition still stands. Further analysis with markers of neuronal differentiation is required to determine whether removing Plzfa and/or Plzfb elicits a phenotype. One way to doing this would be to carry out global expression analysis by RNA-Seq of the mutant compared to wild type embryos. This was an original aim of generating zebrafish mutant lines, but was abandoned when I was unable recapitulate the same phenotype observed in these morphants.

Prospective transcriptional targets

Unfortunately, due to the failure of the ChIP-Seq experiment I was unable to obtain a list of transcriptional targets that Plzf in order to understand how it is carrying out its function in Chapter 3. In other systems, Plzf is known to directly bind

and either repress or activate transcription as described within Chapter 1. A number of these target genes have roles in the CNS that could account for the function of Plzf in inhibiting neurogenesis.

Examples of this include members of the Id family of proteins, which have roles in inhibiting regulating proliferation and differentiation of various cell types, including in the nervous system. There is evidence that Plzf is able to drive expression of *Id2* in both myeloid and natural killer T cells (Doulatov et al., 2009, Gleimer et al., 2012). In the spermatogonial stem cells, Plzf positively regulates expression of *Redd1,* which controls the ability of these cells to self-renew (Hobbs et al., 2010). This same gene has been investigated in the developing mouse cerebral cortex where they have shown Redd1 acts to prevent cell cycle exit and neuronal differentiation (Malagelada et al., 2011).

Potential redundancy with other family members

Studies in the mouse and chicken spinal cord have demonstrated a function for Plzf in progenitor maintenance that isn't redundant to Notch signalling (Gaber et al., 2013). The absence of the same phenotype in the mutant zebrafish hindbrain may highlight differences in the control of neurogenesis between different species or the different tissues. It should be noted that the phenotypes observed in the CNS of other organisms are relatively mild (Gaber et al., 2013) and the adult knockout mice are viable (Barna et al., 2000).

The partial phenotype observed upon removal of Plzf function may be indicative of functional redundancy with other BTB domain containing family members. In the mouse genome, a highly similar gene to *Plzf* known as *Fazf* (also termed *Plzf2* or *Zbtb32*) has been analysed (Zhang et al., 1999). This gene, which is believed to have arisen from a genome duplication event, encodes a protein with a high degree of similarity to Plzf, has been found expressed in a similar pattern to Plzf in hematopoietic progenitor cells and its forced expression is also capable of inhibiting cell proliferation (Dai et al., 2002). Further evidence has shown Fazf is capable of specifically binding the same target sequences as Plzf and that it can heterodimerise with Plzf (Hoatlin et al., 1999). It has been hypothesised that the Fazf protein is functionally redundant to Plzf in the developing hematopoietic

system, accounting for the lack of a relevant phenotype seen in the knockout mouse (Mcconnell and Licht, 2007).

Another interesting candidate for functional redundancy is the Dpzf (Zbtb20) protein. This transcription factor is known to interact with the same DNA binding site as Plzf and is expressed during hippocampal neurogenesis in the mouse (Mitchelmore et al., 2002). Performing a search on the Zebrafish Information Network (Sprague et al., 2006) reveals that at least 4 other *zbtb* family members are expressed in the developing zebrafish CNS (*zbtb2b, zbtb4, zbtb8os, zbtb10*). Whether these or other transcription factors have functional roles in the developing CNS and can compensate for the absence of Plzfa remains to be studied.

Consistent with the idea that Plzf is redundant with other factors, there is evidence from previous work that Plzfa is not the only target of Btbd6a for degradation during zebrafish primary neurogenesis (Sobieszczuk et al., 2010). Ectopic expression of Btbd6a results in a dramatic increase in the amount of cells undergoing neurogenesis, yet knockdown of Plzfa does not reveal the identical phenotype.

Plzfa acts at multiple steps in the neurogenic cascade

In this thesis I present evidence that Plzf is expressed in cells that are undergoing neuronal differentiation as well as in neural progenitors. Both the *plzfa:eGFP* transient transgenic embryos and analysis by fluorescent ISH suggest that *plzfa* transcription is increased in cells exiting the ventricular zone and beginning to undergo differentiation. This upregulation coincides with expression of the proneural gene *neurog1*, which is therefore a candidate transcription factor that is responsible for *plzfa* expression. I attempted to confirm this using gain and loss-of-function experiments, but was unable to determine whether *neurog1* is upstream of *plzfa* transcription. Interestingly, these cells expressing higher levels of the transcript did not appear to contain higher levels of Plzfa protein.

The previously published results showed that Plzfa is able to inhibit expression of *neurog1* (Sobieszczuk et al., 2010) and therefore prevent the onset of neuronal differentiation. The model proposes that differentiating cells are able to overcome this inhibition because of the Neurog1-diven expression of *btbd6a* that

leads to degradation of Plzfa. The observation that endogenous Plzf protein is present in cells beginning to undergo neurogenesis suggests that Plzfa protein is removed at a later stage in the neurogenic cascade and poses the question of whether it has a function in these cells.

As well as my expression analysis, previous observations also provided some evidence that Plzfa functions at later stages of differentiation. Partial knockdown of Btbd6a in zebrafish resulted in a decrease of late markers of neurogenesis, such as *deltaB* and *islet1*, but had little effect on *neurog1* (Sobieszczuk et al., 2010). In contrast, complete knockdown of Btbd6a led to a decrease of all markers of neurogenesis. Studies on the *Xenopus* homologue of *btbd6* found that knockdown results in decreased expression of late markers without affecting early markers (Bury et al., 2008). Finally, *btbd6* was found to be a direct transcriptional target of NeuroD but not Neurog1 in *Xenopus* ectodermal explants (Seo et al., 2007). These results suggest that *btbd6a* expression and function may also be acting during late stages of neuronal differentiation.

Other genes that have a similar expression pattern to *plzfa* are known to regulate genes expressed at different phases of the neurogenic cascade. One example is miR-9, which has an expression pattern that is reminiscent to *plzfa* at 42 hpf in the zebrafish hindbrain (Coolen et al., 2012). miR-9 functions both in progenitors, having an important role relating to progenitor maintenance (Bonev et al., 2011, Bonev et al., 2012), and also in differentiating cells where it prevents cell cycle exit by inhibiting *elavl3 (HuC)* transcription (Coolen et al., 2012). I therefore aimed to determine whether Plzfa was acting at multiple steps in the neurogenic cascade.

To test this, I aimed to force cells to undergo neurogenesis and determine whether the addition of Plzfa would inhibit expression of late neurogenic markers, such as *neurod4*. Injecting embryos with *neurog1* RNA resulted in ectopic *neurod4* expressing cells, whilst embryos coinjected with Plzfa RNA showed wild-type levels of *neurod4* expression. This is consistent with the idea that Plzfa is able to inhibit neurogenesis by reducing expression of proneural genes at multiple steps in the cascade. Knowledge of the direct transcriptional targets would provide us with the mechanism how Plzf is able to do this.

Although the experiments are not yet complete and final conclusions cannot be drawn, I can put forward a hypothesis to account for the results. In neural progenitor cells in the ventricular zone, Neurog1 protein expression is either at low levels or is oscillating, depending upon whether the oscillatory dynamics observed in the neural epithelium of other vertebrates holds true for the zebrafish. In these progenitor cells, an unknown factor is driving *plzfa* transcription and the Plzfa protein is acting to inhibit *neurog1* expression. The decision to commit towards a neurogenic fate is largely governed by Notch-mediated lateral inhibition.

When cells begin to exit the ventricular zone and initiate the neurogenic cascade, they express sustained high levels of *neurog1*. Neurog1 drives transcription of other proneural genes, such as *neurod4* (Seo et al., 2007), therefore beginning the neurogenic cascade, and also activates expression of *btbd6a*. The previous model suggested that this would lead to degradation of Plzfa in these cells and allow neurogenesis to continue (Sobieszczuk et al., 2010). Based upon the expression patterns I observe, I propose a different model where Neurog1 also leads to increased expression of *plzfa* transcript in these cells. The simultaneous activation of *btbd6a* expression ensures sustained levels of Plzfa protein in these cells which restrain both *neurog1* and *neurod4* transcription, although its action on *neurog1* is insufficient to prevent its high levels of transcription. The ability of Plzfa to keep levels of *neurod4* low could account for why I only observe minimal overlap between *neurog1* and *neurod4* in cells undergoing neurogenesis. This is described diagrammatically in Figure 7-1b.

Late differentiating cells express high levels of *neurod4*, and low levels of *neurog1* and *plzfa*. Genes downstream of the Neurogenins have been shown to indirectly inhibit expression of *neurogenin* (Liu and Harland, 2005) and therefore it is likely that NeuroD4 is able to carry out a similar role. Plzfa's ability to inhibit the proneural genes is attenuated by the presence of Btbd6a acting to remove Plzfa from the cells and thus allow *neurod4* expression. The absence of *plzfa* transcription in these cells could be explained by inhibition caused by NeuroD4, along with, or separate to, a loss of upstream activators of *plzfa* expression. This allows the neurogenic cascade to continue and therefore neuronal differentiation, shown in Figure 7-1c.

Figure 7-1: Model of Plzfa action

Hypothesis of the role of Plzfa at multiple steps of the neurogenic cascade.

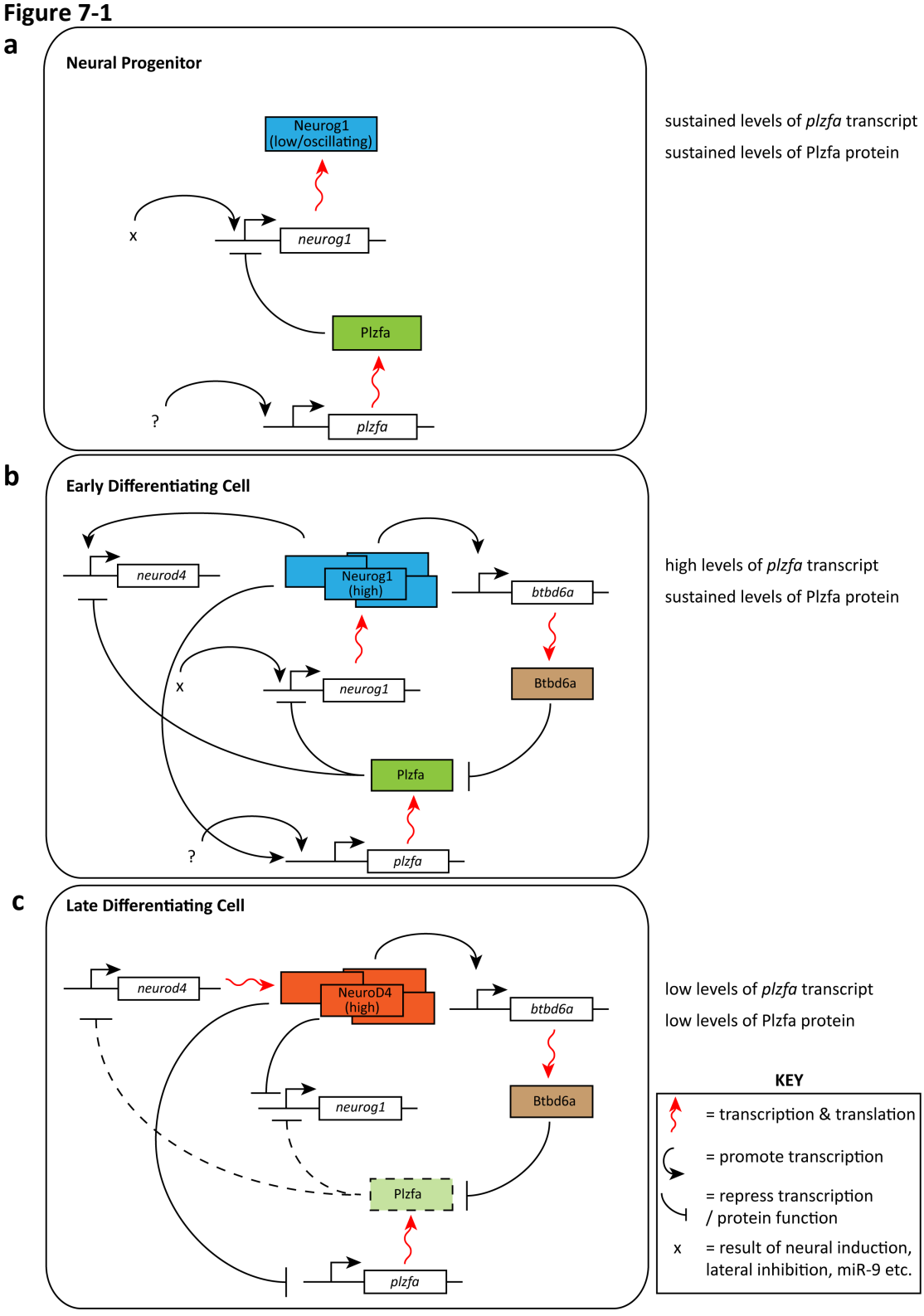
a: Neural progenitors contain low or oscillating Neuog1. This is as a result of various inputs to proneural gene expression described within Chapter 1 and simplified here to 'x'. *plzfa* transcripts and protein are found in these cells, driven by unknown '?' factor(s), and function to keep *neurog1* levels low.

b: Cells beginning to differentiate contain high levels of Neurog1. This drives expression of *btbd6a,* which leads to attenuation of Plzfa protein expression and activity. High levels of Neurog1 also increase transcription of *plzfa,* accounting for the higher levels of transcript in those cells. This ensures that some Plzfa protein remains and this functions to inhibit expression of both *neurog1* and *neurod4*. Neurog1 is known to drive transcription of *neurod4*, but Plzfa inhibits its expression and therefore keeps the levels low in these early differentiating cells.

c: In late differentiating cells NeuroD4 is expressed at high levels. The absence of *plzfa* transcription in these cells suggests that NeuroD4 is able to lead to the inhibition of *plzfa* expression. Published evidence that NeuroD4 is able to drive expression of *btbd6a* suggests that the adaptor protein is able to remove Plzfa protein from these cells, ensuring that there is no protein left (indicated by dashed lines) to inhibit proneural gene expression. There is also published evidence that NeuroD4 is able to indirectly inhibit *neurog1* expression, helping explain why cells containing high levels of *neurod4* contain low or no *neurog1* expression.

The arrows indicate proposed regulatory relationships, which may be direct or indirect.

Figure 7-1


a
Neural Progenitor
Neurog1 (low/oscillating)
x
neurog1
Plzfa
?
plzfa
sustained levels of plzfa transcript
sustained levels of Plzfa protein
b
Early Differentiating Cell
neurod4
Neurog1 (high)
btbd6a
x
neurog1
Btbd6a
Plzfa
?
plzfa
high levels of plzfa transcript
sustained levels of Plzfa protein
c
Late Differentiating Cell
neurod4
NeuroD4 (high)
btbd6a
neurog1
Btbd6a
Plzfa
plzfa
low levels of plzfa transcript
low levels of Plzfa protein
KEY
= transcription & translation
= promote transcription
= repress transcription / protein function
x = result of neural induction, lateral inhibition, miR-9 etc.

I have presented this model in order to explain the expression patterns that I observe in the zebrafish hindbrain, taking into account the neurogenic function of Plzfa revealed by overexpression experiments in this work and from Sobieszczuk et al. (2010). This model predicts that the function of Plzfa is to ensure cells undergo the complete neurogenic cascade and do not prematurely upregulate late markers of neurogenesis. Further functional analysis is required to determine if the hypothesised processes are happening during zebrafish neurogenesis. One prediction from this model that can be directly tested is that removing Plzfa function should result in a *neurod4* being expressed prematurely. This may result in increased overlap of expression between *neurog1* and *neurod4* in the differentiating cells.

It should also be noted that the majority of my expression analysis was carried out at 44 hpf, whereas the work by Sobieszczuk et al. (2010) focused on the 12 hpf neuroepithelium. It is therefore possible that the differences in expression of *plzfa* RNA and protein that I observe compared to the published work could be accounted for by differences in mechanisms underlying neurogenesis at these different stages.

Genome modification technology

In Chapter 5 I show that TALEN technology can be used to generate loss-of-function mutations in target genes. I show that these TALENs can be highly efficient at inducing mutations specific to the target site and screening for these mutations can be done by a number of methods, including HRM analysis. Until recently, morpholino-mediated gene knockdown technology constituted the main method by which zebrafish researchers could manipulate endogenous gene function. As discussed, the off-target effects caused by many morpholinos require that phenotypes related to gene knockdown to be confirmed by other methods, ideally by functional analysis of embryos containing mutations in the same gene.

I therefore anticipate that widespread use of TALENs, along with the more recently developed CRISPR/Cas9 technology, will significantly improve our ability to understand gene function in zebrafish. Along with this work, published work (Law and Sargent, 2014) and unpublished observations from the 11th International

Conference on Zebrafish Development and Genetics (Madison, USA), researchers have demonstrated that some mutants are unable to phenocopy morphants for the same gene. In some cases this is likely to reflect off-target effects related to morpholino usage, however care must also be taken to ensure mutants are truly abolishing protein function. The random nature of indel formation means that some mutant alleles won't lead to frame shifts that disrupt protein function. Recent work has shown that by injecting a stop codon cassette for insertion by homology-dependent repair ensures open reading frame truncation (Gagnon et al., 2014). As discussed, the use of double CRISPR to remove large regions of gDNA can be used to eliminate entire exons and ensure protein function is lost.

I devised a simple technique for the insertion of exogenous DNA into targeted locations of the zebrafish genome aided by the use of TALENs within Chapter 6. This method does not rely on long homology arms and can be easily adapted to work with a TALEN pair of one's choosing. Furthermore, the insertion event appeared specific to the targeted location. Through the insertion of eGFP and GAL4 I show that I obtain a high frequency of integration into the *plzfa* locus in injected embryos and by employing various selection methods I obtain a germline transmission efficiency of around 25%. The expression pattern observed in the transgenic embryos appeared to match broadly with the expected *plzfa* pattern.

This method is, however, not without its drawbacks. In particular, the ligation event was associated with the insertion and deletion of nucleotides. Due to the way in which the insertion technique was designed, this frequently results the cassette being inserted out-of-frame with the targeted protein, therefore preventing the cassette being correctly translated. This imprecision has also been reported when using similar techniques in certain cultured cells (Maresca et al., 2013) and in zebrafish (Auer et al., 2014). Furthermore, similar imprecision has been associated with insertions through injection of DNA containing short homology arms into zebrafish (Bedell et al., 2012, Hwang et al., 2013a), but not when injecting into mouse embryos (Yang et al., 2013). A picture is therefore emerging that species-specific effects influence the precision of the insertion events, although further analysis is required. As discussed in Chapter 6, the technique can be adapted such that the imprecision does not impact on the translation of the cassette.

This method adds to our growing repertoire of techniques to specifically manipulate the zebrafish genome. The technique's main advantages over the previously published homology dependent insertion techniques that make use of TALENs are the improved germline transmission, the ability to insert very large cassettes and the simplicity of inserting different TALEN binding sites and cassettes into the plasmid. Figure 7-2 outlines the main features of the methods available for insertions using TALENs at the time of writing.

In this work, I have demonstrated the application of the technique to make use of the *plzfa* transcriptional machinery to drive fluorescent proteins and the transcriptional activator GAL4. This provides us with the ability to monitor gene expression in living embryos and to drive expression of transgenes in a spatially restricted manner. Commandeering the endogenous promoter and enhancers to drive transgenes has various advantages for functional analysis compared to techniques such as RNA injection, which results in transient overexpression of the construct throughout the embryo. This technique could therefore be used to express an epitope tagged form of Plzf, for example for ChIP-Seq analysis, or to drive expression of dominant negative versions of the transcription factor, such as those used for dissecting Plzf function in the spinal cord (Gaber et al., 2013).

Spatial control over transgene expression could also be combined with temporal control. Techniques such as heat shock (Esengil and Chen, 2008) or drug-inducible GAL4 expression (Gerety et al., 2013) can be used to specifically switch on expression at the desired point in time. These methods could be used to scrutinise Plzf function at later stages in neurogenesis, such as in differentiating cells migrating down the glial fibres, without affecting primary neurogenesis.

For many years, the ability to specifically add recombinase sites at targeted locations in the mouse genome has meant that researchers have been able to temporally and spatially switch off gene function (Metzger and Chambon, 2001). Gene targeting technology in the zebrafish brings us into an era where generating conditional knockouts is possible. This will allow us to analyse gene function in cases where loss of the protein is lethal to the embryo during early development and define functions for genes in specific tissues without affecting the result of the organism.

Figure 7-2: Comparison of TALEN-mediated insertion methods in zebrafish

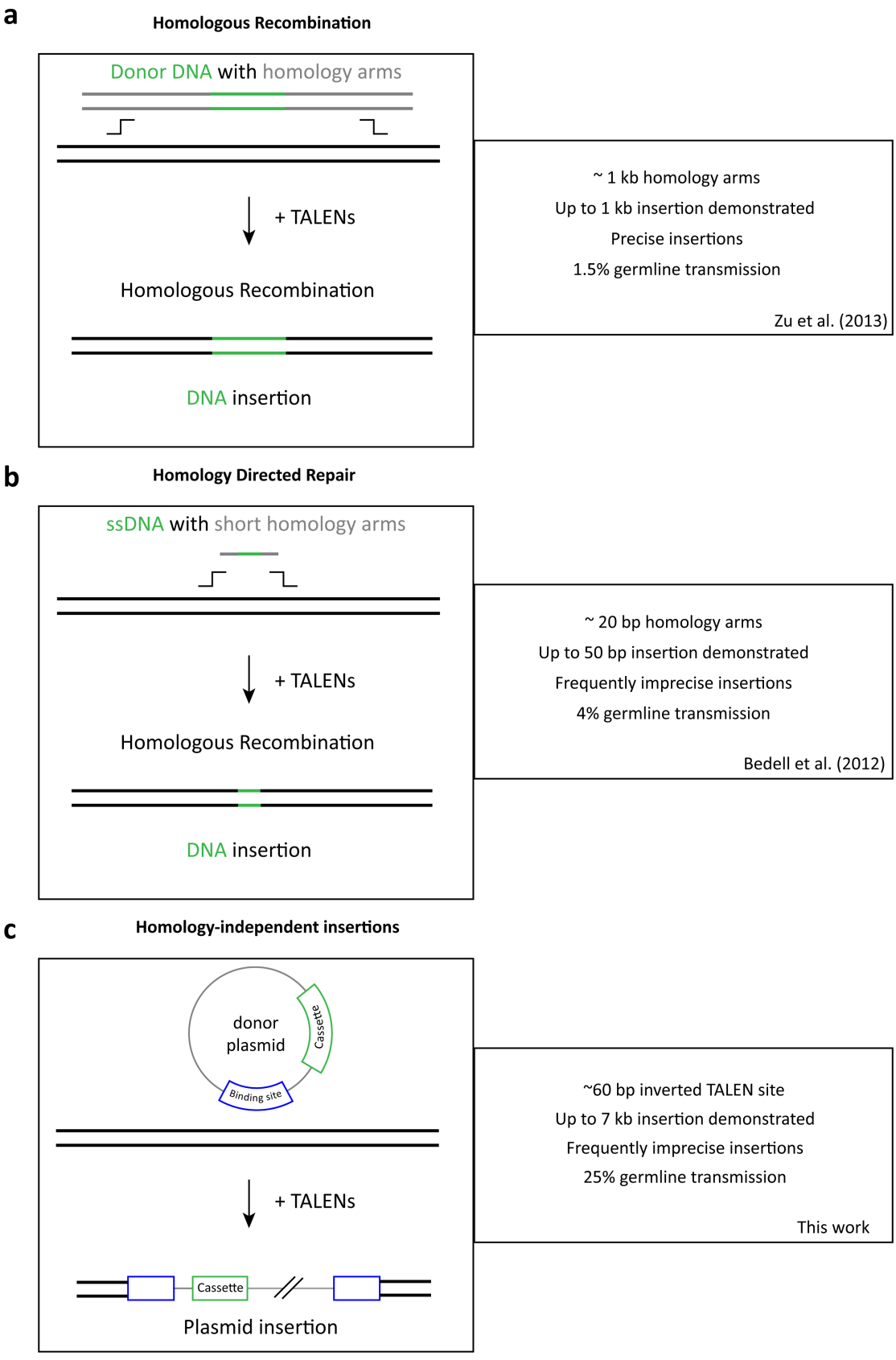
Figures depicting the 3 reported methods for targeted insertions in the zebrafish genome using TALENs.

a: Homologous recombination has been demonstrated in the zebrafish by Zu et al. (2013). Using homology arms of around 1 kb each they insert an eGFP cassette just under 1 kb into a specific loci in the zebrafish genome. They find that 4/275 (1.5%) injected fish transmit the insertion and all of them do so precisely.

b: Homology directed repair was shown by Bedell et al. (2012). They are able to insert up to 50 bp of DNA flanked by 20 bp homology arms by injecting ssDNA along with the TALENs. They observe integration associated with indels in the majority of cases and find that 8/186 (4%) transmit the insertion.

c: The homology independent insertion that I have described in this thesis. Instead of providing homology arms, I place the inverted TALEN target sequence upstream of the insertion cassette. I have demonstrated the ability to insert up to 7 kb plasmids using this technique, however the insertion event is associated with indels. Finally, I find that 7/27 (26%) of injected embryos selected for the insertion transmit to their progeny.

Figure 7-2


a
Homologous Recombination
Donor DNA with homology arms
+ TALENs
Homologous Recombination
DNA insertion
~ 1 kb homology arms
Up to 1 kb insertion demonstrated
Precise insertions
1.5% germline transmission
Zu et al. (2013)
b
Homology Directed Repair
ssDNA with short homology arms
+ TALENs
Homologous Recombination
DNA insertion
~ 20 bp homology arms
Up to 50 bp insertion demonstrated
Frequently imprecise insertions
4% germline transmission
Bedell et al. (2012)
c
Homology-independent insertions
donor plasmid
Cassette
Binding site
+ TALENs
Cassette
Plasmid insertion
~60 bp inverted TALEN site
Up to 7 kb insertion demonstrated
Frequently imprecise insertions
25% germline transmission
This work

Final comments

An important method to understand protein function is to investigate the phenotype observed upon disruption of the gene. Conventional forward genetic approaches that rely on random mutagenesis are well established for the zebrafish and have been used to generate numerous mutant alleles. A downside of this method is that it is expensive, time- and labour-intensive and doesn't allow the targeting of specific genes. Therefore the advent of MO technology to specifically and rapidly knockdown gene function has been widely adopted for analysis in various organisms, including zebrafish (Summerton and Weller, 1997). Reports regarding the off-target activation of pro-apoptotic genes by MOs have revealed the requirement for careful controls, including blocking Tp53 function, in order to account for artefacts associated with MO use.

As demonstrated in this work, the establishment of gene targeting methods such as TALENs and CRISPR/Cas9 result in the rapid generation of mutant alleles. This means that that MO-induced phenotypes are starting to be tested by generating these mutants. This work shows that the phenotype observed in embryos injected with *plzfa* and *plzfb* MOs in a *tp53* mutant background is not consistent with the double mutant zebrafish. In agreement with recently published work (Law and Sargent, 2014), this suggests that MOs can lead to spurious phenotypes irrespective of Tp53 function. It therefore seems highly likely that gene targeting technology will supplant MO usage for analysis of gene function. This statement echoes a recent opinion paper that suggests future descriptions of phenotypes based solely on MOs would be judged very critically without comparison to a genetic mutant (Schulte-Merker and Stainier, 2014).

A large amount of research has gone into understanding the tight control regarding the onset of neurogenesis, in particular the various inputs into Notch-mediated lateral inhibition. The expression of high levels of Neurog1 is known to be sufficient to initiate a cascade of downstream proneural genes, yet we are lacking knowledge regarding control over these subsequent steps of neuronal differentiation. In this work, I present evidence that Plzfa is expressed at multiple steps during neuronal differentiation and hypothesise that it is important to ensure

the correct timing of proneural gene expression during the pathway to differentiation.

8 Appendix

Figure 8-1: Nucleotide identity between *plzfa* and *plzfb*

a: Pairwise sequence alignment of *plzfa* (first line) and *plzfb* (third line) cDNA, starting from the start codon. Identical (|) nucleotides are marked on the second line. The alignment was carried out using EMBOSS needle (http://www.ebi.ac.uk/Tools/psa/emboss_needle) using the Needleman-Wunsch algorithm. Overall identity: 74.8% (1564/2092) and gaps 9.0% (188/2092). Regions marked in blue correspond to the forward and reverse primers used to generate the *plzfa* ISH probe and regions in red mark the *plzfb* probe. Identity between the probe regions drops to 61.4% (584/951).

b: Comparison of the sequence immediately upstream of the translation start sites for *plzfa* and *plzfb*. Kozak consensus sequences are highlighted in green and the start codon is bolded. The second Kozak sequence present in the *plzfb* UTR would result in an out-of-frame protein being produced.

Figure 8-1

a


plzfa 1 ATGGATTTGACTAAAATGGGAATGATCCAGCTCCAGAACCCCAATCACCC 50
plzfb 1 ATGGATTTGACTAAAATGGGTGCGATCCACCTTCAGAACCCCAGCCACCC 50
plzfa 51 --CAATGCCCTGCTGCACAAGGCTAATCAGATGCGCCTGGCTGGCACGCT 98
plzfb 51 TACAA--CTCTCCTGCAGAAGGCCAACCAGATGCGCCTGTCTGGGACATT 98
plzfa 99 GTGCGACGTGGTCATCATGGTGGACAGCCAGGAGTTCCACGCTCACCGCA 148
plzfb 99 GTGTGATGTGGTCATCATGGTCGACAGCCAGGAGTTTCATGCGCACAGAA 148
plzfa 149 CCGTGCTGGCCTGCACCAGTAAGATGTTCGAAATCCTCTTTCACCGGAAC 198
plzfb 149 CCGTGCTTGCATGTGCTAGTAAGATGTTCGAGATCCTGTTCCACCGCAAC 198
plzfa 199 AGCCAACATTACACTCTTGACTTTCTCTCACCAAAGACCTTCCAGCAGAT 248
plzfb 199 AGCCAACATTACACTCTGGACTTCCTCTCACCAAAGACCTTCCAGCAGAT 248
plzfa 249 TCTGGAGTATGCCTACACTGCCACGCTCCAAGCTAAAGTAGAGGATCTGG 298
plzfb 249 TTTGGAGTACGCCTACACCGCCACGCTCCAAGCCAAGCTGGAAGACCTGG 298
plzfa 299 ATGATCTGCTGTACGCAGCAGAGATCTTAGAAATAGAGTACTTGGAGGAG 348
plzfb 299 ATGACCTGCTGTACGCCGCAGAGATTTTAGAAATCGAGTACTTGGAGGAG 348
plzfa 349 CAATGCCTGAAGATTCTGGAGACCATCCAGTCCTCGGATGAAAATGACAC 398
plzfb 349 CAATGTCTGAAGATTTTGGAAACCATCCAGGCATCTGAAGAGAACGATAC 398
plzfa probe forward primer
plzfa 399 TGAGGTGAACATGAACGATGGTGGCACAGAA-GATGACGAGGA-GCCGGAA 446
plzfb 399 AGATGTGAACATGAATG----------AGAATGGTGCCGAGGATGTTGAA 438
plzfa 447 GGGACGGCACGGA--AGGAACCTGGTCGGT---------TCGAAAAAGCA 485
plzfb 439 --GACCGCA-GGATCAAGTACGTTAAGAGTGCATTTATCTCTAAAAAGCA 485
plzfb probe forward primer
plzfa 486 TTCCACAGAAGAGAGCGGCTACGTGTCA--GCTGCCCAGCAAGCACTTGC 533
plzfb 486 TTCCATACAAGATGGCGGTTA--TGGCACCGCTGCCC-GCCATCGCCTG- 531
plzfa 534 GCTGCCAGGT----ATGGTGGATCAGAGTCCCTCAGTATCCACTTCC--- 576
plzfb 532 GCT--CTGGTCAGCATGGTCGATAAAAGATCTTCAGAGT---CTGCCCAA 576
plzfa 577 TTCGGCCTCTCCACCATGAGCCCAACCAAAGCAGCCGTGGACAGCCTAAT 626
plzfb 577 CTGGGTCTTTCAACTGTCAGCCCTACTAAAGCTGCAGTAGAGGGCTTGAT 626
plzfa 627 GAGCATTGGTCAGTCCCTCCTGCAGAGCACCATGCACCCTGGCGTGGGTG 676
plzfb 627 GAGCATGGGACAGTCCGTCCTCCAAGGCAACAT----------------- 659
plzfa 677 CTGAACAG---CCCTTGCA------TGGCAACTCGCA-CCCCA-TGATGG 715
plzfb 660 -TGGACAGAGTTTCTCGCACAACATTGGTAACT--CATCTCCATTGAT-A 705
plzfa 716 GAGAA-ATCAAAACGGAGATGATGCAAGTAGATGAGAGCGGA----GAAC 760
plzfb 706 GATAAGATAAAAACTGAGATGATGCAAGTAGACGA----GGACTATCAAC 751
plzfa 761 ATGAGAGCCCTAAAGCCATGGAGTCAATCGCCTCCAGCAATGGTGAGCGC 810
plzfb 752 ATGAAGG-----------TGGCTTC-----CCT--AGGAACGGGGAGGGT 783
plzfa 811 AGCGGAGA-GCCAGACAAGAACCGTGATGGGCCGGGTACCCCTACCAGGA 859
plzfb 784 GGCGGTGACTCCAGGC-AGCACCGGGACACTCCTGGAACCCCGACCAGAG 832
plzfa 860 GCAGCGTGATCACCAGTGCCCGCGAGCTGCACTACGTCAGGGATGAAGGT 909
plzfb 833 GCAGCGTCATCACCAGCGCCCGCGAGCTCCACT---CCACCGAT--AGTT 877
plzfa 910 CTGGGTGA-TCAGCAGGCCGAGG-----TCAGCCAAATGGGGCTGGAGGC 953
plzfb 878 CT-GTTGACTC------CCAAGGGGACAACAACCAA-----------GC 908
plzfa 954 C------------ATGGCGGGAATGACTGAAAAACATTTGGCGTCACTCT 991
plzfb 909 CAGTTTGATCAGTATGGCCAGTCTGGCTGATAGACATTTCGCTTCGCTGT 958
plzfa 992 ACGGCATTCCTTCTAATCACAAAAATG-AAGCGATGCT-CTCGATGCCTG 1039
plzfb 959 GCTCCATCCCGTCCAACCATAAGAGCGAAAGC-ATGCTGCCCG-TGCCTG 1006
plzfa 1040 CCTCCATGGCCTCATCCTTGCACATGT--CCCCAGCCCTGGCCATGTCAA 1087
plzfb 1007 TGTCCATGGCCTCGTCCCT--ACATGTGCCGCCATCCCTGGCCATGTCTA 1054

b

plzfa *5′* ATGAGTTGGCCGTGTGGATTTCAGAGACGTT**ATGG**ATTGA *3′*

plzfb *5′* TCTGTGAGATGGTTCGGTAAATCTTGATGTT**ATGG**ATTGA *3′*

Figure 8-2: Protein similarity between Plzfa and Plzfb

Alignment of the Plzfa (first line) and Plzfb proteins (third line). Identical (|) and similar (:) amino acids are marked on the second line. The most N-terminal blue shaded region indicates the amino acids present in the BTB domain (Ahmad et al., 1998) and the second blue region indicates the RD2 domain located between amino acids 200-300 (Li et al., 1997). The regions shaded in red indicate each of the 9 zinc fingers. The alignment was done using EMBOSS needle (http://www.ebi.ac.uk/Tools/psa/emboss_needle) using the Needleman-Wunsch algorithm. The C2H2 Zinc Finger motifs were determined using SMART (http://smart.embl-heidelberg.de/). Overall identity: 75.3% (508/675), similarity: 84.1% (568/675) and gaps: 3.0% (20/675).

Figure 8-2

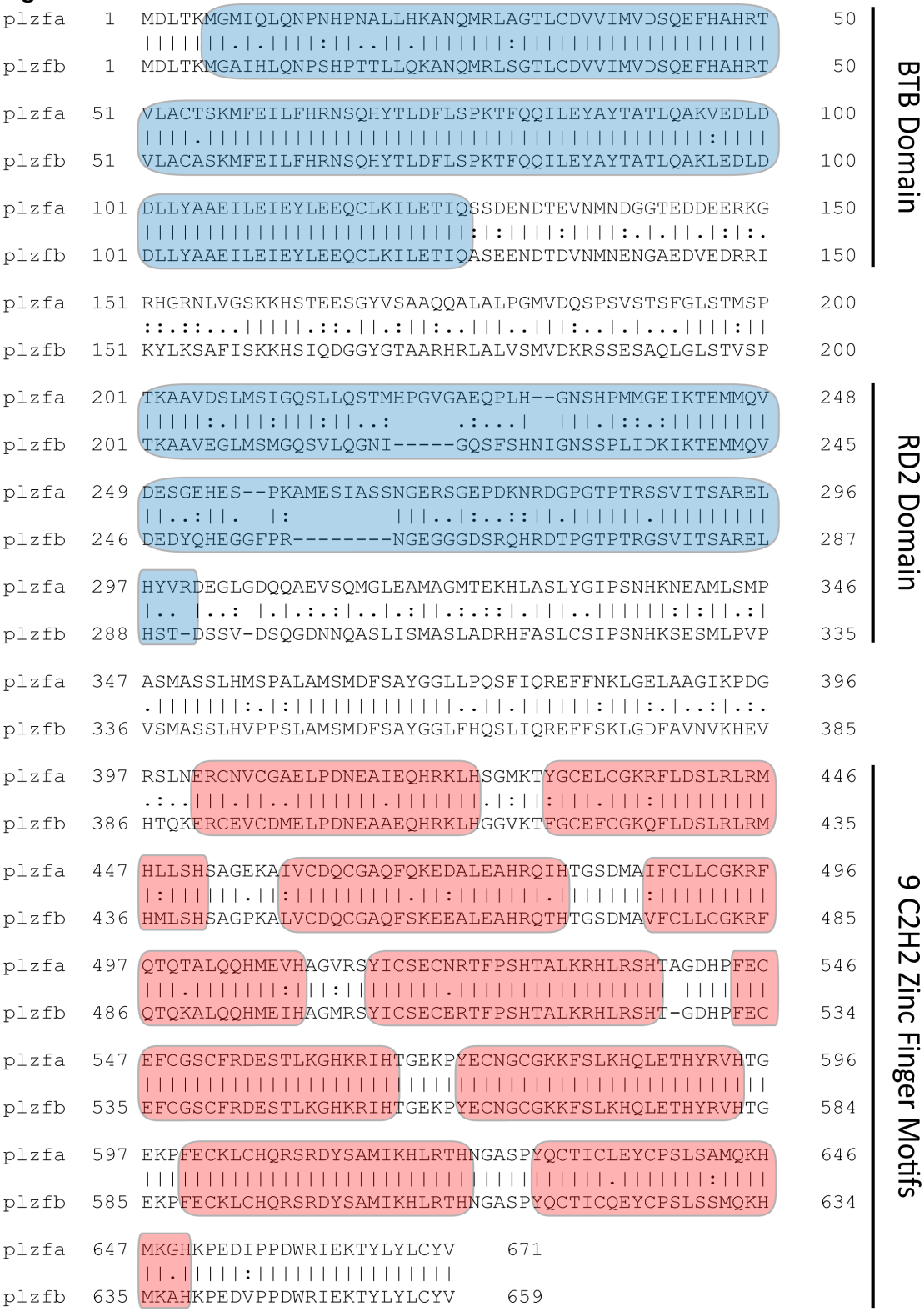

plzfa 1 MDLTKMGMIQLQNPNHPNALLHKANQMRLAGTLCDVVIMVDSQEFHAHRT 50
plzfb 1 MDLTKMGAIHLQNPSHPTTLLQKANQMRLSGTLCDVVIMVDSQEFHAHRT 50
plzfa 51 VLACTSKMFEILFHRNSQHYTLDFLSPKTFQQILEYAYTATLQAKVEDLD 100
plzfb 51 VLACASKMFEILFHRNSQHYTLDFLSPKTFQQILEYAYTATLQAKLEDLD 100
plzfa 101 DLLYAAEILEIEYLEEQCLKILETIQSSDENDTEVNMNDGGTEDDEERKG 150
plzfb 101 DLLYAAEILEIEYLEEQCLKILETIQASEENDTDVNMNENGAEDVEDRRI 150
plzfa 151 RHGRNLVGSKKHSTEESGYVSAAQQALALPGMVDQSPSVSTSFGLSTMSP 200
plzfb 151 KYLKSAFISKKHSIQDGGYGTAARHRLALVSMVDKRSSESAQLGLSTVSP 200
plzfa 201 TKAAVDSLMSIGQSLLQSTMHPGVGAEQPLH--GNSHPMMGEIKTEMMQV 248
plzfb 201 TKAAVEGLMSMGQSVLQGNI-----GQSFSHNIGNSSPLIDKIKTEMMQV 245
plzfa 249 DESGEHES--PKAMESIASSNGERSGEPDKNRDGPGTPTRSSVITSAREL 296
plzfb 246 DEDYQHEGGFPR--------NGEGGGDSRQHRDTPGTPTRGSVITSAREL 287
plzfa 297 HYVRDEGLGDQQAEVSQMGLEAMAGMTEKHLASLYGIPSNHKNEAMLSMP 346
plzfb 288 HST-DSSV-DSQGDNNQASLISMASLADRHFASLCSIPSNHKSESMLPVP 335
plzfa 347 ASMASSLHMSPALAMSMDFSAYGGLLPQSFIQREFFNKLGELAAGIKPDG 396
plzfb 336 VSMASSLHVPPSLAMSMDFSAYGGLFHQSLIQREFFSKLGDFAVNVKHEV 385
plzfa 397 RSLNERCNVCGAELPDNEAIEQHRKLHSGMKTYGCELCGKRFLDSLRLRM 446
plzfb 386 HTQKERCEVCDMELPDNEAAEQHRKLHGGVKTFGCEFCGKQFLDSLRLRM 435
plzfa 447 HLLSHSAGEKAIVCDQCGAQFQKEDALEAHRQIHTGSDMAIFCLLCGKRF 496
plzfb 436 HMLSHSAGPKALVCDQCGAQFSKEEALEAHRQTHTGSDMAVFCLLCGKRF 485
plzfa 497 QTQTALQQHMEVHAGVRSYICSECNRTFPSHTALKRHLRSHTAGDHPFEC 546
plzfb 486 QTQKALQQHMEIHAGMRSYICSECERTFPSHTALKRHLRSHT-GDHPFEC 534
plzfa 547 EFCGSCFRDESTLKGHKRIHTGEKPYECNGCGKKFSLKHQLETHYRVHTG 596
plzfb 535 EFCGSCFRDESTLKGHKRIHTGEKPYECNGCGKKFSLKHQLETHYRVHTG 584
plzfa 597 EKPFECKLCHQRSRDYSAMIKHLRTHNGASPYQCTICLEYCPSLSAMQKH 646
plzfb 585 EKPFECKLCHQRSRDYSAMIKHLRTHNGASPYQCTICQEYCPSLSSMQKH 634
plzfa 647 MKGHKPEDIPPDWRIEKTYLYLCYV 671
plzfb 635 MKAHKPEDVPPDWRIEKTYLYLCYV 659
BTB Domain
RD2 Domain
9 C2H2 Zinc Finger Motifs

Figure 8-3: *plzfa* MO induces a developmental delay

The extent of developmental delay caused by *plzfa* MO was monitored using the *cldnb:lyngfp* transgenic line. The primordium begins migrating at around 20 hpf and travels towards to the posterior of the embryo, depositing neuromasts on the way (Ghysen and Dambly-Chaudière, 2007). In control MO injected embryos (a), the primordium had migrated to somite 23 at 36 hpf (b) (8/8). In *plzfa* MO injected embryos (c), at the same time point the primordium had only migrated to somite 21 (d) (6/8) – indicating a delay in development. Leaving the embryos to develop for a further 4 hours at 28.5°C (e) allowed the primordium of the *plzfa* morphant embryos to reach somite 23 (f). No delay was observed in embryos injected with *plzfb* MO (5/5). *plzfa* was not observed to be expressed in the lateral line (*data not shown*).

Scale bar = 100 µm

Figure 8-3

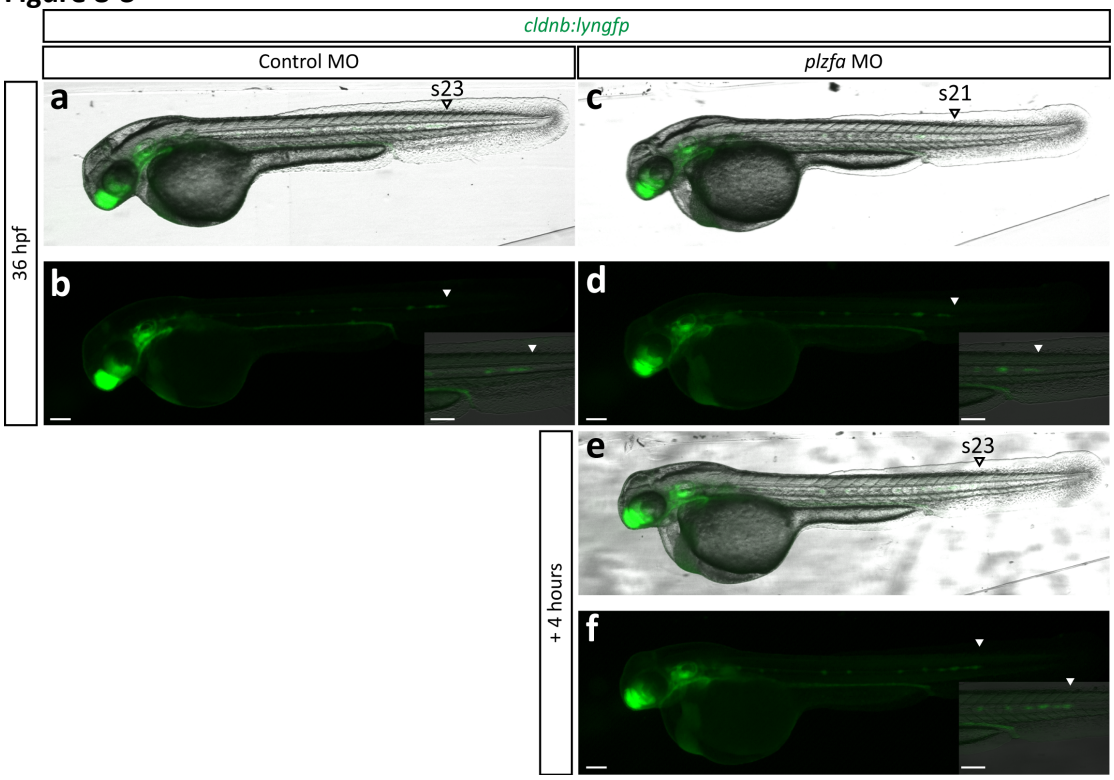

cldnb:lyngfp
Control MO
plzfa MO
36 hpf
+ 4 hours
a
b
c
d
e
f
s23
s21
s23

Figure 8-4: Phenotype is observed in time and stage matched embryos

a-f: Confocal slices of control MO (a), *plzfa* MO (b), *plzfb* MO (c) and double morphant (d) embryos at 44 hpf showing HuC/D. Embryos injected with *plzfa* MO were allowed to develop for a further 4 hours (e & f) until they reached the same stage as control MO injected embryos.

g-i: Time course of HuC/D expression. Dorsal view confocal slices of wild type embryos at 30 (g), 40 (h) and 48 (i) hpf showing stereotypical arrangement of HuC/D cells is maintained throughout the stages shown. Embryos at the indicated stages are shown in a sagittal view, stained for HuC/D (green), zfPlzfa (red) and GFAP (blue) (j-l).

Scale bar = 20 µm

Figure 8-4

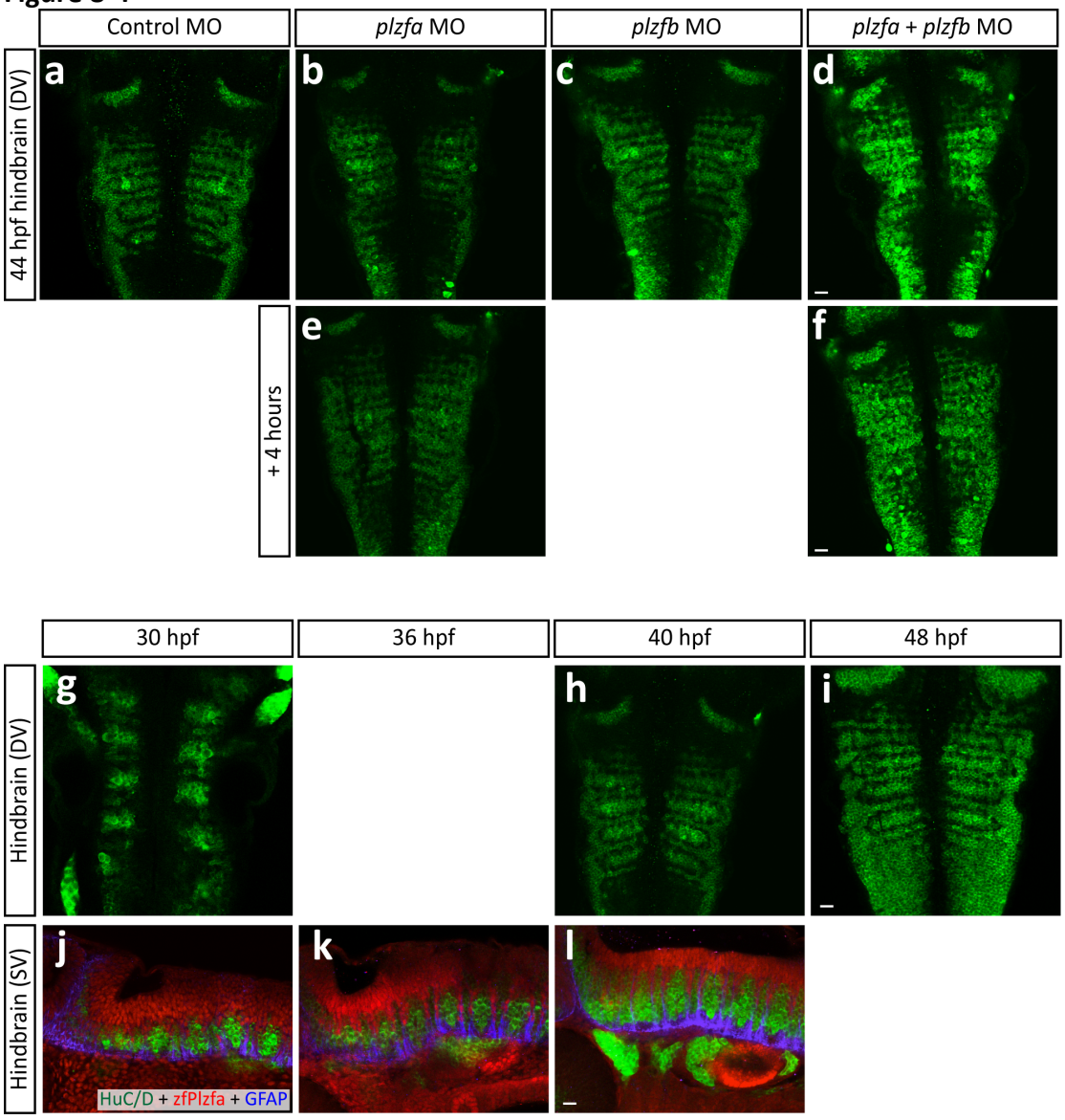

Control MO
plzfa MO
plzfb MO
plzfa + *plzfb* MO
44 hpf hindbrain (DV)
a
b
c
d
+ 4 hours
e
f
30 hpf
36 hpf
40 hpf
48 hpf
Hindbrain (DV)
g
h
i
Hindbrain (SV)
j
k
l
HuC/D + zfPlzfa + GFAP

Figure 8-5: *redd1* and *btbd6a* TALENs

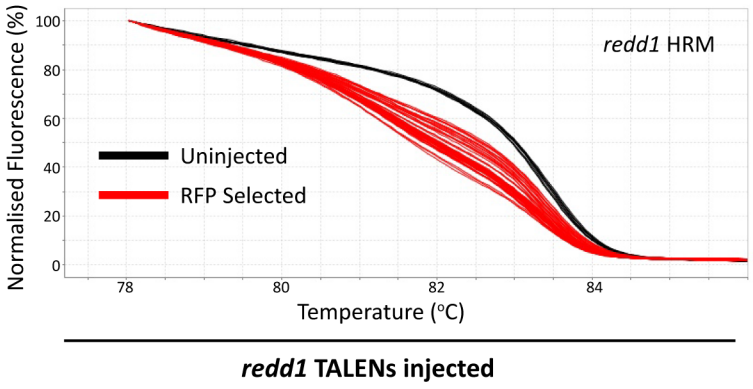
a: HRM analysis for embryos injected with *redd1* TALENs. Melt profiles of injected embryos differ from that of uninjected embryos, indicating the presence of mutations at the *redd1* locus.

b: Injected embryos developed until adulthood with no obvious morphological defects. They were then outcrossed to wild type and the embryos produced were analysed by HRM. Shown here are 3 different genotypes from the same founder fish.

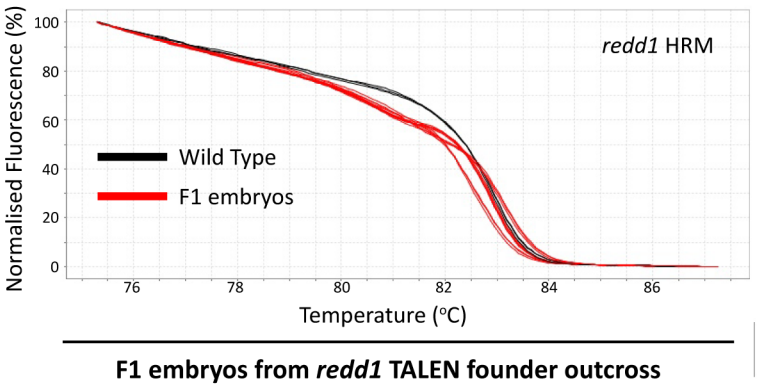
c: Embryos injected with *btbd6a* TALENs show evidence of indels at the *btbd6a* locus when analysed by HRM.

Figure 8-5

a



b



c

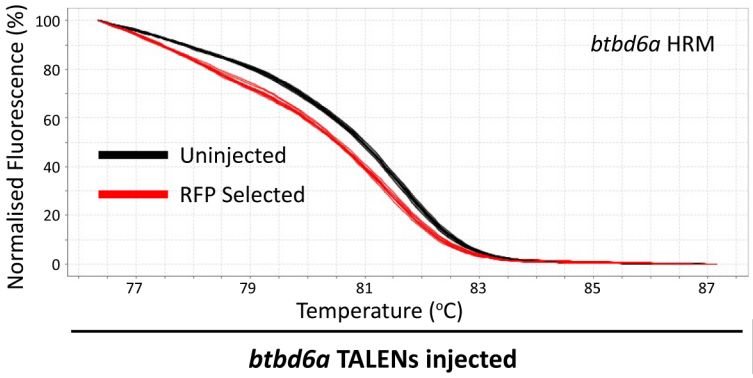


Figure 8-6: *plzfa* donor plasmid sequence

a: Map of the *plzfa* Donor Plasmid.

b: Sequence of the inverted *plzfa* TALEN binding site and *attP* flanked P2A-eGFP insertion cassette. Restriction enzymes shown can be used to replace different modules in the plasmid.

Figure 8-6

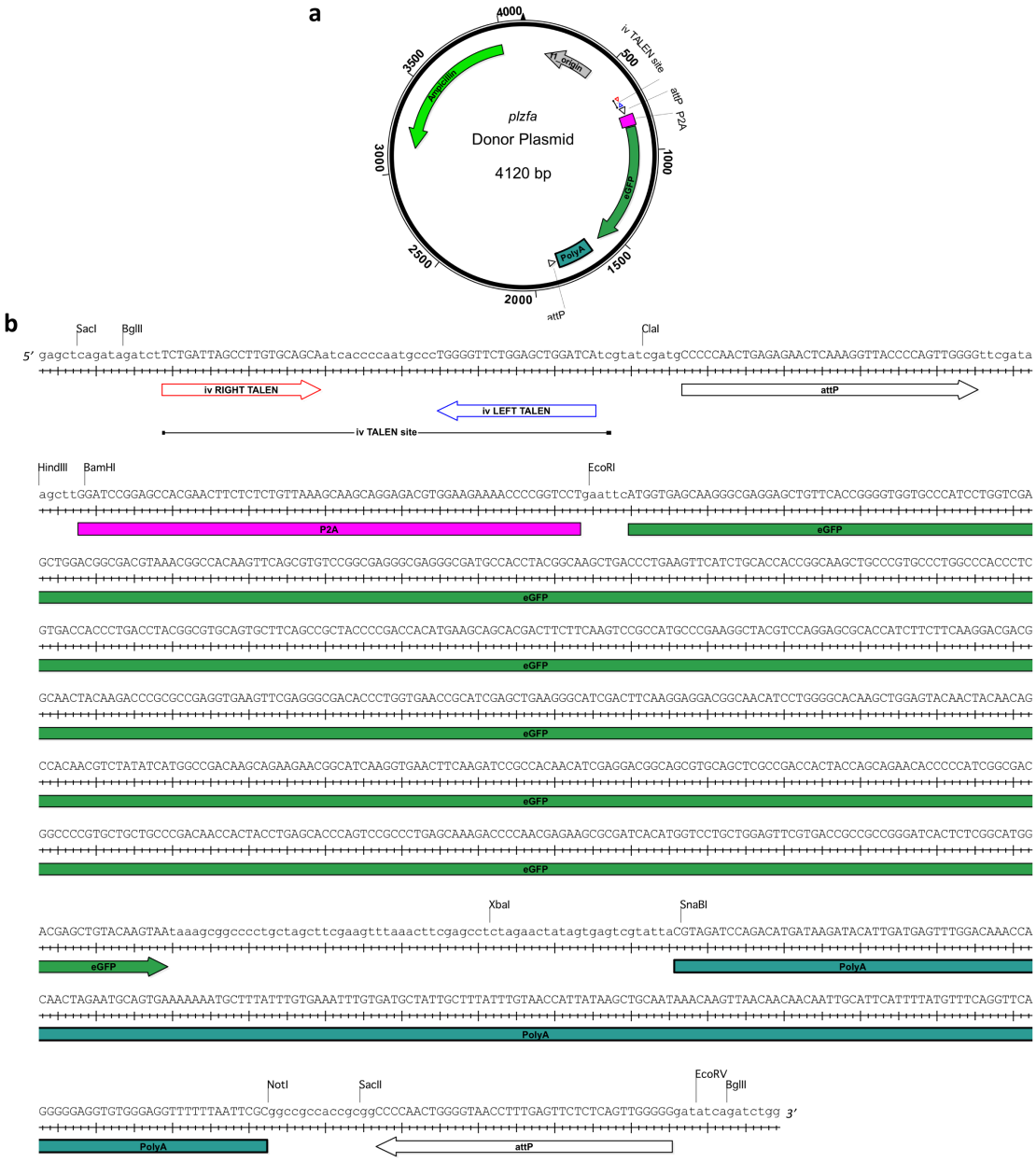

a
4000
500
1000
1500
2000
2500
3000
3500
iv TALEN site
attP
P2A
Ampicillin
Ar origin
plzfa
Donor Plasmid
4120 bp
eGFP
PolyA
attP
b
SacI
BglII
ClaI
5′ gagctcagatagatctTCTGATTAGCCTTGTGCAGCAatcaccccaatgcccTGGGGTTCTGGAGCTGGATCAtcgtatcgatgCCCCCAACTGAGAGAACTCAAAGGTTACCCCAGTTGGGGttcgata
iv RIGHT TALEN
attP
iv LEFT TALEN
iv TALEN site
HindIII
BamHI
EcoRI
agcttGGATCCGGAGCCACGAACTTCTCTCTGTTAAAGCAAGCAGGAGACGTGGAAGAAAACCCCGGTCCTgaattcATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGA
P2A
eGFP
GCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTC
eGFP
GTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACG
eGFP
GCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAG
eGFP
CCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGAC
eGFP
GGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG
eGFP
XbaI
SnaBI
ACGAGCTGTACAAGTAAtaaagcggcccctgctagcttcgaagtttaaacttcgagcctctagaactatagtgagtcgtattaCGTAGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCA
eGFP
PolyA
CAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCA
PolyA
NotI
SacII
EcoRV
BglII
GGGGGAGGTGTGGGAGGTTTTTTAATTCGCggccgccaccgcggCCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGGgatatcagatctgg 3′
PolyA
attP

Figure 8-7: Insertion at the *golden* locus results in imprecise ligation

a-d: The inverted TALEN target sequence in the *plzfa* donor plasmid was replaced with the inverted sequences for the *golden* TALEN pair (a). The *golden* donor plasmid and *golden* TALENs were injected into zebrafish embryo and GFP expression was observed in the embryos (*data not shown*). The expected ligation product is shown (b). The *5'* integration site was amplified using the primers shown (F3 & R1) and a band indicating integration was observed in the expected embryos (c). These bands were excised and sequenced, revealing a range of different insertions and deletions around the integration site (d). Some of these resulted in eGFP being produced in-frame as indicated.

Figure 8-7

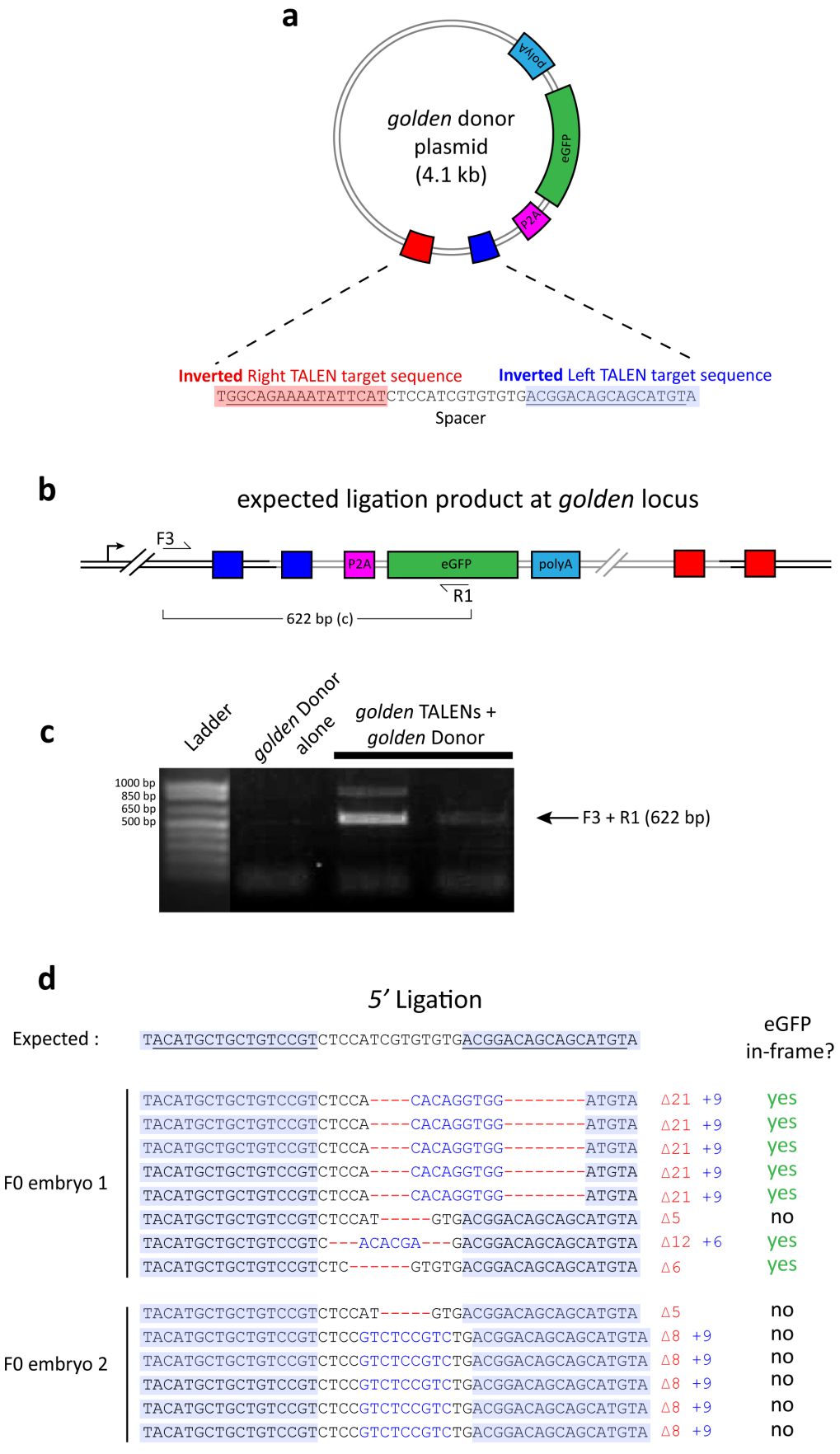

a
golden donor plasmid (4.1 kb)
polyA
eGFP
P2A
Inverted Right TALEN target sequence
Inverted Left TALEN target sequence
TGGCAGAAAATATTCATCTCCATCGTGTGTGACGGACAGCAGCATGTA
Spacer
b
expected ligation product at golden locus
F3
P2A
eGFP
polyA
R1
622 bp (c)
c
Ladder
golden Donor alone
golden TALENs + golden Donor
1000 bp
850 bp
650 bp
500 bp
F3 + R1 (622 bp)
d
5' Ligation
Expected :
TACATGCTGCTGTCCGTCTCCATCGTGTGTGACGGACAGCAGCATGTA
eGFP in-frame?
F0 embryo 1
TACATGCTGCTGTCCGTCTCCA----CACAGGTGG--------ATGTA Δ21 +9 yes
TACATGCTGCTGTCCGTCTCCA----CACAGGTGG--------ATGTA Δ21 +9 yes
TACATGCTGCTGTCCGTCTCCA----CACAGGTGG--------ATGTA Δ21 +9 yes
TACATGCTGCTGTCCGTCTCCA----CACAGGTGG--------ATGTA Δ21 +9 yes
TACATGCTGCTGTCCGTCTCCA----CACAGGTGG--------ATGTA Δ21 +9 yes
TACATGCTGCTGTCCGTCTCCAT-----GTGACGGACAGCAGCATGTA Δ5 no
TACATGCTGCTGTCCGTC---ACACGA---GACGGACAGCAGCATGTA Δ12 +6 yes
TACATGCTGCTGTCCGTCTC------GTGTGACGGACAGCAGCATGTA Δ6 yes
F0 embryo 2
TACATGCTGCTGTCCGTCTCCAT-----GTGACGGACAGCAGCATGTA Δ5 no
TACATGCTGCTGTCCGTCTCCGTCTCCGTCTGACGGACAGCAGCATGTA Δ8 +9 no
TACATGCTGCTGTCCGTCTCCGTCTCCGTCTGACGGACAGCAGCATGTA Δ8 +9 no
TACATGCTGCTGTCCGTCTCCGTCTCCGTCTGACGGACAGCAGCATGTA Δ8 +9 no
TACATGCTGCTGTCCGTCTCCGTCTCCGTCTGACGGACAGCAGCATGTA Δ8 +9 no
TACATGCTGCTGTCCGTCTCCGTCTCCGTCTGACGGACAGCAGCATGTA Δ8 +9 no

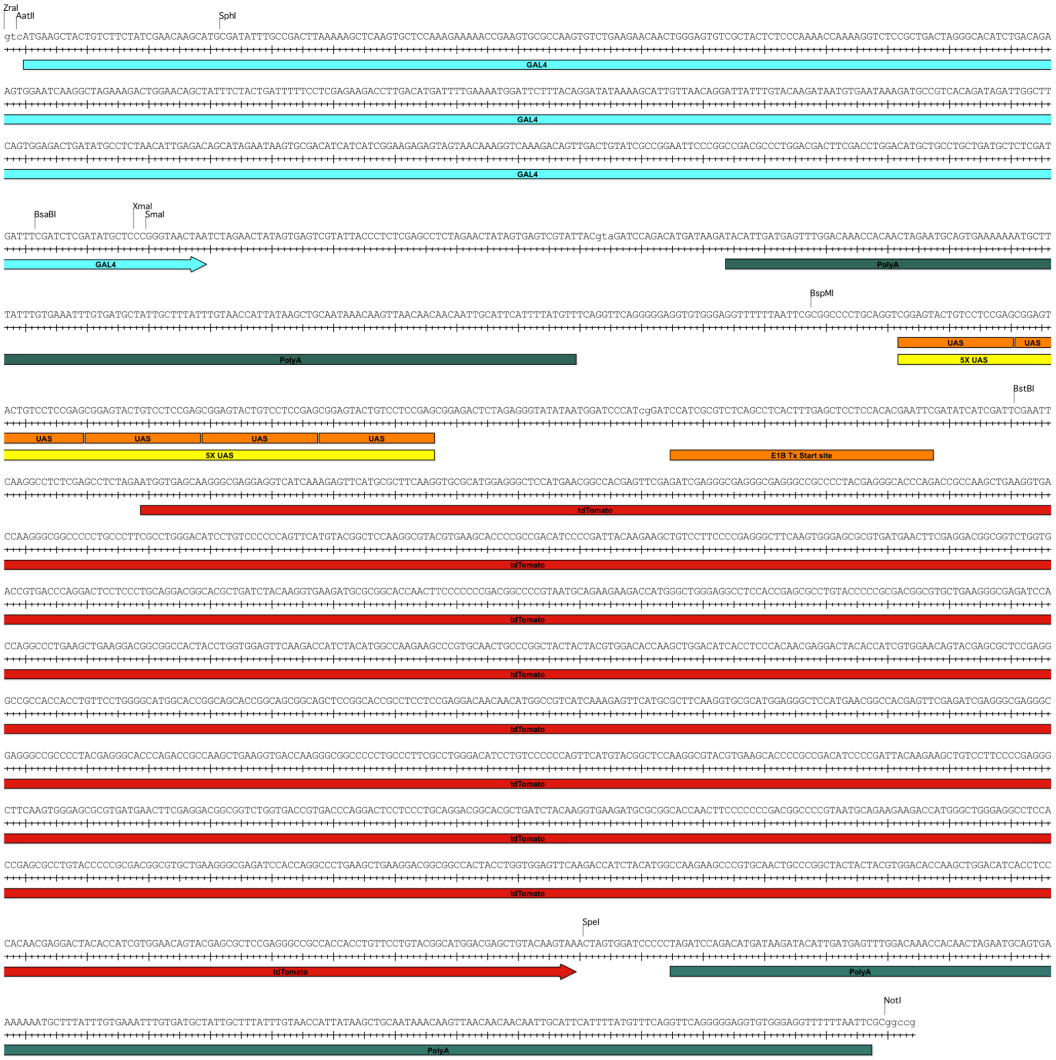
Figure 8-8: *GAL4; UAS:tdTomato* cassette and internal reporter sequences

a: Sequence for the *GAL4; UAS:tdTomato* cassette. *tdTomato* transcription is under control of 5 GAL4 binding sites (*UAS*).

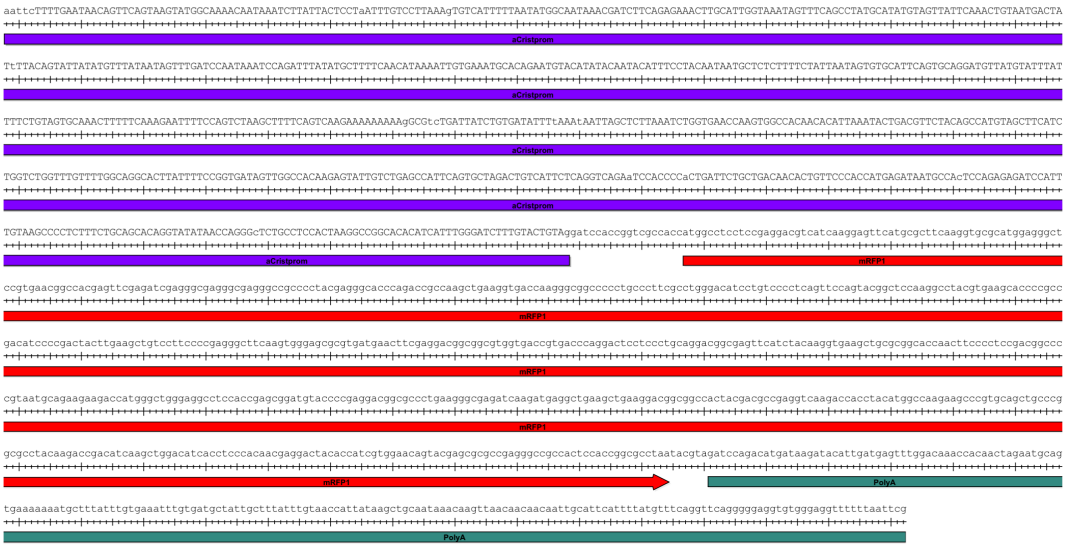
b: Sequence for the *alpha crystalline: mRFP* (ACR) internal reporter.

Figure 8-8

a


ZraI
AatII
SphI
GAL4
BsaBI
XmaI
SmaI
PolyA
BspMI
UAS
5X UAS
BstBI
E1B Tx Start site
tdTomato
SpeI
NotI

b


aCristprom
mRFP1
PolyA

Bibliography

Aday, A. W., Zhu, L. J., Lakshmanan, A., Wang, J. & Lawson, N. D. (2011). Identification of cis regulatory features in the embryonic zebrafish genome through large-scale profiling of H3K4me1 and H3K4me3 binding sites. *Dev. Biol.,* **357**, 450-462.

Adolf, B., Chapouton, P., Lam, C. S., Topp, S., Tannhäuser, B., Strähle, U., Götz, M. & Bally-Cuif, L. (2006). Conserved and acquired features of adult neurogenesis in the zebrafish telencephalon. *Dev. Biol.,* **295**, 278-93.

Ahmad, K. F., Engel, C. K. & Privé, G. G. (1998). Crystal structure of the BTB domain from PLZF. *Proc. Natl. Acad. Sci. U. S. A.,* **95**, 12123-8.

Alexandre, P., Reugels, A. M., Barker, D., Blanc, E. & Clarke, J. D. W. (2010). Neurons derive from the more apical daughter in asymmetric divisions in the zebrafish neural tube. *Nat. Neurosci.,* **13**, 673-9.

Ali, F., Hindley, C., Mcdowell, G., Deibler, R., Jones, A., Kirschner, M., Guillemot, F. & Philpott, A. (2011). Cell cycle-regulated multi-site phosphorylation of Neurogenin 2 coordinates cell cycling with differentiation during neurogenesis. *Development,* **138**, 4267-4277.

Altmann, C. R. & Brivanlou, A. H. (2001). Neural patterning in the vertebrate embryo. *Int. Rev. Cytol.,* **203**, 447-82.

Alvarez-Buylla, A., García-Verdugo, J. M. & Tramontin, A. D. (2001). A unified hypothesis on the lineage of neural stem cells. *Nat. Rev. Neurosci.,* **2**, 287-93.

Amoyel, M., Cheng, Y.-C., Jiang, Y.-J. & Wilkinson, D. G. (2005). Wnt1 regulates neurogenesis and mediates lateral inhibition of boundary cell specification in the zebrafish hindbrain. *Development,* **132**, 775-85.

Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. (1999). Notch signaling: cell fate control and signal integration in development. *Science,* **284**, 770-6.

Auer, T. O., Duroure, K., De Cian, A., Concordet, J.-P. & Del Bene, F. (2014). Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res.,* **24**, 142-53.

Avantaggiato, V., Pandolfi, P. P., Ruthardt, M., Hawe, N., Acampora, D., Pelicci, P. G. & Simeone, A. (1995). Developmental analysis of murine Promyelocyte Leukemia Zinc Finger (PLZF) gene expression: implications for the neuromeric model of the forebrain organization. *J. Neurosci.,* **15**, 4927-42.

Bae, S., Bessho, Y., Hojo, M. & Kageyama, R. (2000). The bHLH gene Hes6, an inhibitor of Hes1, promotes neuronal differentiation. *Development,* **127**, 2933-43.

Baena-Lopez, L. A., Alexandre, C., Mitchell, A., Pasakarnis, L. & Vincent, J.-P. (2013). Accelerated homologous recombination and subsequent genome modification in Drosophila. *Development,* **140**, 4818-25.

Balciunas, D., Wangensteen, K. J., Wilber, A., Bell, J., Geurts, A., Sivasubbu, S., Wang, X., Hackett, P. B., Largaespada, D. A., Mcivor, R. S. & Ekker, S. C. (2006). Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. *PLoS Genet.,* **2**, e169-e169.

Ball, H. J., Melnick, A., Shaknovich, R., Kohanski, R. A. & Licht, J. D. (1999). The promyelocytic leukemia zinc finger (PLZF) protein binds DNA in a high molecular weight complex associated with cdc2 kinase. *Nucleic Acids Res.,* **27**, 4106-13.

Bally-Cuif, L. & Hammerschmidt, M. (2003). Induction and patterning of neuronal development, and its connection to cell cycle control. *Curr. Opin. Neurobiol.,* **13**, 16-25.

Bardwell, V. J. & Treisman, R. (1994). The POZ domain: a conserved protein-protein interaction motif. *Genes Dev.,* **8**, 1664-1677.

Barna, M., Hawe, N., Niswander, L. & Pandolfi, P. P. (2000). Plzf regulates limb and axial skeletal patterning. *Nat. Genet.,* **25**, 166-72.

Barna, M., Merghoub, T., Costoya, J. A., Ruggero, D., Branford, M., Bergia, A., Samori, B. & Pandolfi, P. P. (2002). PLZF mediates transcriptional repression of HoxD Gene Expression through Chromatin Remodeling. *Dev. Cell,* **3**, 499-510.

Becker, C. G. & Becker, T. (2008). Adult zebrafish as a model for successful central nervous system regeneration. *Restor. Neurol. Neurosci.,* **26**, 71-80.

Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug, R. G., Tan, W., Penheiter, S. G., Ma, A. C., Leung, A. Y. H., Fahrenkrug, S. C., Carlson, D. F., Voytas, D. F., Clark, K. J., Essner, J. J. & Ekker, S. C. (2012). In vivo genome editing using a high-efficiency TALEN system. *Nature,* **491**, 114-8.

Bedell, V. M., Westcot, S. E. & Ekker, S. C. (2011). Lessons from morpholino-based screening in zebrafish. *Briefings in functional genomics,* **10**, 181-8.

Berghmans, S., Murphey, R. D., Wienholds, E., Neuberg, D., Kutok, J. L., Fletcher, C. D. M., Morris, J. P., Liu, T. X., Schulte-Merker, S., Kanki, J. P., Plasterk, R., Zon, L. I. & Look, A. T. (2005). Tp53 Mutant Zebrafish Develop Malignant Peripheral Nerve Sheath Tumors. *Proc. Natl. Acad. Sci. U. S. A.,* **102**, 407-12.

Bernardo, M. V., Yelo, E., Gimeno, L., Campillo, J. A. & Parrado, A. (2007). Identification of apoptosis-related PLZF target genes. *Biochem. Biophys. Res. Commun.,* **359**, 317-22.

Bernardos, R. L. & Raymond, P. A. (2006). GFAP transgenic zebrafish. *Gene expression patterns : GEP,* **6**, 1007-13.

Bertrand, N., Castro, D. S. & Guillemot, F. (2002). Proneural genes and the specification of neural cell types. *Nat. Rev. Neurosci.,* **3**, 517-30.

Bill, B. R., Petzold, A. M., Clark, K. J., Schimmenti, L. A. & Ekker, S. C. (2009). A primer for morpholino use in zebrafish. *Zebrafish,* **6**, 69-77.

Bitinaite, J., Wah, D. A., Aggarwal, A. K. & Schildkraut, I. (1998). FokI dimerization is required for DNA cleavage. *Proc. Natl. Acad. Sci. U. S. A.,* **95**, 10570-10575.

Blader, P. & Strähle, U. (2000). Zebrafish developmental genetics and central nervous system development. *Hum. Mol. Genet.,* **9**, 945-51.

Blankenberg, D., Von Kuster, G., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A. & Taylor, J. (2010). Galaxy: a web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol,* **Chapter 19**, Unit 19.10.1-21.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. & Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science,* **326**, 1509-12.

Bogdanove, A. J. & Voytas, D. F. (2011). TAL effectors: customizable proteins for DNA targeting. *Science,* **333**, 1843-6.

Bogdanović, O., Fernández-Miñán, A., Tena, J. J., De La Calle-Mustienes, E. & Gómez-Skarmeta, J. L. (2013). The developmental epigenomics toolbox: ChIP-seq and MethylCap-seq profiling of early zebrafish embryos. *Methods,* **62**, 207-15.

Bond, A. M., Bhalala, O. G. & Kessler, J. A. (2012). The dynamic role of bone morphogenetic proteins in neural stem cell fate and maturation. *Dev. Neurobiol.,* **72**, 1068-84.

Bonev, B., Pisco, A. & Papalopulu, N. (2011). MicroRNA-9 reveals regional diversity of neural progenitors along the anterior-posterior axis. *Dev. Cell,* **20**, 19-32.

Bonev, B., Stanley, P. & Papalopulu, N. (2012). MicroRNA-9 Modulates Hes1 ultradian oscillations by forming a double-negative feedback loop. *Cell reports,* **2**, 10-8.

Bonn, S., Zinzen, R. P., Perez-Gonzalez, A., Riddell, A., Gavin, A.-C. & Furlong, E. E. M. (2012). Cell type-specific chromatin immunoprecipitation from multicellular complex samples using BiTS-ChIP. *Nat. Protoc.,* **7**, 978-94.

Borello, U., Cobos, I., Long, J. E., Mcwhirter, J. R., Murre, C. & Rubenstein, J. L. R. (2008). FGF15 promotes neurogenesis and opposes FGF8 function during neocortical development. *Neural Dev.,* **3**, 17-17.

Boukarabila, H., Saurin, A. J., Batsché, E., Mossadegh, N., Van Lohuizen, M., Otte, A. P., Pradel, J., Muchardt, C., Sieweke, M. & Duprez, E. (2009). The PRC1 Polycomb group complex interacts with PLZF/RARA to mediate leukemic transformation. *Genes Dev.,* **23**, 1195-206.

Bourguignon, C., Li, J. & Papalopulu, N. (1998). XBF-1, a winged helix transcription factor with dual activity, has a role in positioning neurogenesis in Xenopus competent ectoderm. *Development,* **125**, 4889-900.

Brand, A. H. & Livesey, F. J. (2011). Neural stem cell biology in vertebrates and invertebrates: more alike than different? *Neuron,* **70**, 719-29.

Bray, S. J. (2006). Notch signalling: a simple pathway becomes complex. *Nat. Rev. Mol. Cell Biol.,* **7**, 678-89.

Breau, M. A., Wilson, D., Wilkinson, D. G. & Xu, Q. (2012). Chemokine and Fgf signalling act as opposing guidance cues in formation of the lateral line primordium. *Development,* **139**, 2246-53.

Brewster, R., Lee, J. & Ruiz I Altaba, A. (1998). Gli/Zic factors pattern the neural plate by defining domains of cell differentiation. *Nature,* **393**, 579-83.

Brunet, J. F. & Ghysen, A. (1999). Deconstructing cell determination: proneural genes and neuronal identity. *Bioessays,* **21**, 313-8.

Buaas, F. W., Kirsh, A. L., Sharma, M., Mclean, D. J., Morris, J. L., Griswold, M. D., De Rooij, D. G. & Braun, R. E. (2004). Plzf is required in adult male germ cells for stem cell self-renewal. *Nat. Genet.,* **36**, 647-52.

Bultje, R. S., Castaneda-Castellanos, D. R., Jan, L. Y., Jan, Y.-N., Kriegstein, A. R. & Shi, S.-H. (2009). Mammalian Par3 regulates progenitor cell asymmetric division via notch signaling in the developing neocortex. *Neuron,* **63**, 189-202.

Bultmann, S., Morbitzer, R., Schmidt, C. S., Thanisch, K., Spada, F., Elsaesser, J., Lahaye, T. & Leonhardt, H. (2012). Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res.,* **40**, 5368-77.

Bury, F. J., Moers, V., Yan, J., Souopgui, J., Quan, X.-J., De Geest, N., Kricha, S., Hassan, B. A. & Bellefroid, E. J. (2008). Xenopus BTBD6 and its Drosophila homologue lute are required for neuronal development. *Dev. Dyn.,* **237**, 3352-60.

Cade, L., Reyon, D., Hwang, W. Y., Tsai, S. Q., Patel, S., Khayter, C., Joung, J. K., Sander, J. D., Peterson, R. T. & Yeh, J.-R. J. (2012). Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res.*, 1-10.

Capecchi, M. R. (2005). Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat. Rev. Genet.,* **6**, 507-12.

Carroll, D. (2013). Genome Engineering with Targetable Nucleases. *Annu. Rev. Biochem.,* **83**, 140307200228009-140307200228009.

Castro, D. S., Martynoga, B., Parras, C., Ramesh, V., Pacary, E., Johnston, C., Drechsel, D., Lebel-Potter, M., Garcia, L. G., Hunt, C., Dolle, D., Bithell, A., Ettwiller, L., Buckley, N. & Guillemot, F. (2011). A novel function of the proneural factor Ascl1 in progenitor proliferation identified by genome-wide characterization of its targets. *Genes Dev.,* **25**, 930-45.

Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. A., Somia, N. V., Bogdanove, A. J. & Voytas, D. F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.,* **39**, e82-e82.

Chao, T.-T., Chang, C.-C. & Shih, H.-M. (2007). SUMO modification modulates the transrepression activity of PLZF. *Biochem. Biophys. Res. Commun.,* **358**, 475-82.

Chauchereau, A., Mathieu, M., De Saintignon, J., Ferreira, R., Pritchard, L. L., Mishal, Z., Dejean, A. & Harel-Bellan, A. (2004). HDAC4 mediates transcriptional repression by the acute promyelocytic leukaemia-associated protein PLZF. *Oncogene,* **23**, 8777-84.

Chen, H., Thiagalingam, A., Chopra, H., Borges, M. W., Feder, J. N., Nelkin, B. D., Baylin, S. B. & Ball, D. W. (1997). Conservation of the Drosophila lateral inhibition pathway in human lung cancer: a hairy-related protein (HES-1) directly represses achaete-scute homolog-1 expression. *Proc. Natl. Acad. Sci. U. S. A.,* **94**, 5355-60.

Chen, Z., Brand, N. J., Chen, A., Chen, S. J., Tong, J. H., Wang, Z. Y., Waxman, S. & Zelent, A. (1993). Fusion between a novel Krüppel-like zinc finger gene and the retinoic acid receptor-alpha locus due to a variant t(11;17) translocation associated with acute promyelocytic leukaemia. *EMBO J.,* **12**, 1161-7.

Cheng, Y.-C., Amoyel, M., Qiu, X., Jiang, Y.-J., Xu, Q. & Wilkinson, D. G. (2004). Notch activation regulates the segregation and differentiation of rhombomere boundary cells in the zebrafish hindbrain. *Dev. Cell,* **6**, 539-50.

Chitnis, A., Henrique, D., Lewis, J., Ish-Horowicz, D. & Kintner, C. (1995). Primary neurogenesis in Xenopus embryos regulated by a homologue of the Drosophila neurogenic gene Delta. *Nature,* **375**, 761-6.

Chizhikov, V. V. & Millen, K. J. (2004). Mechanisms of roof plate formation in the vertebrate CNS. *Nat. Rev. Neurosci.,* **5**, 808-12.

Chizhikov, V. V. & Millen, K. J. (2005). Roof plate-dependent patterning of the vertebrate dorsal central nervous system. *Dev. Biol.,* **277**, 287-95.

Clarke, J. (2009). Role of polarized cell divisions in zebrafish neural tube formation. *Curr. Opin. Neurobiol.,* **19**, 134-8.

Collins, J. E., White, S., Searle, S. M. J. & Stemple, D. L. (2012). Incorporating RNA-seq data into the zebrafish Ensembl genebuild. *Genome Res.,* **22**, 2067-78.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science,* **339**, 819-23.

Conti, L. & Cattaneo, E. (2010). Neural stem cell systems: physiological players or in vitro entities? *Nat. Rev. Neurosci.,* **11**, 176-87.

Cook, M., Gould, A., Brand, N., Davies, J., Strutt, P., Shaknovich, R., Licht, J., Waxman, S., Chen, Z., Gluecksohn-Waelsch, S., Krumlauf, R. & Zelent, A. (1995). Expression of the Zinc-Finger Gene PLZF at Rhombomere Boundaries in the Vertebrate Hindbrain. *Proc. Natl. Acad. Sci. U. S. A.,* **92**, 2249-2253.

Cooke, J. E., Kemp, H. A. & Moens, C. B. (2005). EphA4 is required for cell adhesion and rhombomere-boundary formation in the zebrafish. *Curr. Biol.,* **15**, 536-42.

Coolen, M., Thieffry, D., Drivenes, Ø., Becker, T. S. & Bally-Cuif, L. (2012). miR-9 Controls the Timing of Neurogenesis through the Direct Inhibition of Antagonistic Factors. *Dev. Cell,* **22**, 1052-1064.

Costoya, J. A., Hobbs, R. M., Barna, M., Cattoretti, G., Manova, K., Sukhwani, M., Orwig, K. E., Wolgemuth, D. J. & Pandolfi, P. P. (2004). Essential role of Plzf in maintenance of spermatogonial stem cells. *Nat. Genet.,* **36**, 653-9.

Costoya, J. A., Hobbs, R. M. & Pandolfi, P. P. (2008). Cyclin-dependent kinase antagonizes promyelocytic leukemia zinc-finger through phosphorylation. *Oncogene,* **27**, 3789-96.

Cunliffe, V. T. (2004). Histone deacetylase 1 is required to repress Notch target gene expression during zebrafish neurogenesis and to maintain the production of motoneurones in response to hedgehog signalling. *Development,* **131**, 2983-95.

Dahlem, T. J., Hoshijima, K., Jurynec, M. J., Gunther, D., Starker, C. G., Locke, A. S., Weis, A. M., Voytas, D. F. & Grunwald, D. J. (2012). Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet.,* **8**, e1002861-e1002861.

Dai, M.-S., Chevallier, N., Stone, S., Heinrich, M. C., Mcconnell, M., Reuter, T., Broxmeyer, H. E., Licht, J. D., Lu, L. & Hoatlin, M. E. (2002). The effects of the Fanconi anemia zinc finger (FAZF) on cell cycle, apoptosis, and proliferation are differentiation stage-specific. *J. Biol. Chem.,* **277**, 26327-34.

David, G., Alland, L., Hong, S. H., Wong, C. W., Depinho, R. A. & Dejean, A. (1998). Histone deacetylase associated with mSin3A mediates repression by the acute promyelocytic leukemia-associated PLZF protein. *Oncogene,* **16**, 2549-56.

Del Bene, F., Wehman, A. M., Link, B. A. & Baier, H. (2008). Regulation of neurogenesis by interkinetic nuclear migration through an apical-basal notch gradient. *Cell,* **134**, 1055-65.

Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J.-K., Shi, Y. & Yan, N. (2012). Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science,* **335**, 720-3.

Dessaud, E., Mcmahon, A. P. & Briscoe, J. (2008). Pattern formation in the vertebrate neural tube: a sonic hedgehog morphogen-regulated transcriptional network. *Development,* **135**, 2489-503.

Dhordain, P., Albagli, O., Honore, N., Guidez, F., Lantoine, D., Schmid, M., The, H. D., Zelent, A. & Koken, M. H. (2000). Colocalization and heteromerization between the two human oncogene POZ/zinc finger proteins, LAZ3 (BCL6) and PLZF. *Oncogene,* **19**, 6240-50.

Diez Del Corral, R., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M. & Storey, K. (2003). Opposing FGF and Retinoid Pathways Control Ventral Neural Pattern, Neuronal Differentiation, and Segmentation during Body Axis Extension. *Neuron,* **40**, 65-79.

Ding, Q., Regan, Stephanie n., Xia, Y., Oostrom, Leoníe a., Cowan, Chad a. & Musunuru, K. (2013). Enhanced Efficiency of Human Pluripotent Stem Cell Genome Editing through Replacing TALENs with CRISPRs. *Cell Stem Cell,* **12**, 393-394.

Distel, M., Wullimann, M. F. & Köster, R. W. (2009). Optimized Gal4 genetics for permanent gene expression mapping in zebrafish. *Proc. Natl. Acad. Sci. U. S. A.,* **106**, 13365-70.

Doetsch, F. (2003). The glial identity of neural stem cells. *Nat. Neurosci.,* **6**, 1127-34.

Dong, S., Zhu, J., Reid, A., Strutt, P., Guidez, F., Zhong, H. J., Wang, Z. Y., Licht, J., Waxman, S. & Chomienne, C. (1996). Amino-terminal protein-protein interaction motif (POZ-domain) is responsible for activities of the promyelocytic leukemia zinc finger-retinoic acid receptor-alpha fusion protein. *Proc. Natl. Acad. Sci. U. S. A.,* **93**, 3624-3624.

Dong, Z., Yang, N., Yeo, S.-Y., Chitnis, A. & Guo, S. (2012). Intralineage Directional Notch Signaling Regulates Self-Renewal and Differentiation of Asymmetrically Dividing Radial Glia. *Neuron,* **74**, 65-78.

Doulatov, S., Notta, F., Rice, K. L., Howell, L., Zelent, A., Licht, J. D. & Dick, J. E. (2009). PLZF is a regulator of homeostatic and cytokine-induced myeloid development. *Genes Dev.,* **23**, 2076-87.

Doyon, Y., Vo, T. D., Mendel, M. C., Greenberg, S. G., Wang, J., Xia, D. F., Miller, J. C., Urnov, F. D., Gregory, P. D. & Holmes, M. C. (2011). Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat. Methods,* **8**, 74-9.

Elkabetz, Y., Panagiotakos, G., Al Shamy, G., Socci, N. D., Tabar, V. & Studer, L. (2008). Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. *Genes Dev.,* **22**, 152-65.

Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. (2009). Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS ONE,* **4**, e5553-e5553.

Esengil, H. & Chen, J. K. (2008). Gene regulation technologies in zebrafish. *Mol. Biosyst.,* **4**, 300-8.

Farah, M. H., Olson, J. M., Sucic, H. B., Hume, R. I., Tapscott, S. J. & Turner, D. L. (2000). Generation of neurons by transient expression of neural bHLH proteins in mammalian cells. *Development,* **127**, 693-702.

Filipponi, D., Hobbs, R. M., Ottolenghi, S., Rossi, P., Jannini, E. A., Pandolfi, P. P. & Dolci, S. (2007). Repression of kit expression by Plzf in germ cells. *Mol. Cell. Biol.,* **27**, 6770-81.

Fischer, S., Kohlhase, J., Böhm, D., Schweiger, B., Hoffmann, D., Heitmann, M., Horsthemke, B. & Wieczorek, D. (2008). Biallelic loss of function of the promyelocytic leukaemia zinc finger (PLZF) gene causes severe skeletal defects and genital hypoplasia. *J. Med. Genet.,* **45**, 731-7.

Fish, J. L., Dehay, C., Kennedy, H. & Huttner, W. B. (2008). Making bigger brains-the evolution of neural-progenitor-cell division. *J. Cell Sci.,* **121**, 2783-93.

Florio, M. & Huttner, W. B. (2014). Neural progenitors, neurogenesis and the evolution of the neocortex. *Development,* **141**, 2182-2194.

Foley, J. E., Yeh, J.-R. J., Maeder, M. L., Reyon, D., Sander, J. D., Peterson, R. T. & Joung, J. K. (2009). Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). *PLoS ONE,* **4**, e4348-e4348.

Ford-Perriss, M., Abud, H. & Murphy, M. (2001). Fibroblast growth factors in the developing central nervous system. *Clin. Exp. Pharmacol. Physiol.,* **28**, 493-503.

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K. & Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.*, 1-6.

Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.,* **32**, 279-84.

Gaber, Z. B., Butler, S. J. & Novitch, B. G. (2013). PLZF regulates fibroblast growth factor responsiveness and maintenance of neural progenitors. *PLoS Biol.,* **11**, e1001676-e1001676.

Gagnon, J. A., Valen, E., Thyme, S. B., Huang, P., Ahkmetova, L., Pauli, A., Montague, T. G., Zimmerman, S., Richter, C. & Schier, A. F. (2014). Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS ONE,* **9**, e98186-e98186.

Gallinari, P., Di Marco, S., Jones, P., Pallaoro, M. & Steinkühler, C. (2007). HDACs, histone deacetylation and gene transcription: from molecular biology to cancer therapeutics. *Cell Res.,* **17**, 195-211.

Garneau, J. E., Dupuis, M.-È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A. H. & Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature,* **468**, 67-71.

Geling, A., Itoh, M., Tallafuss, A., Chapouton, P., Tannhäuser, B., Kuwada, J. Y., Chitnis, A. B. & Bally-Cuif, L. (2003). bHLH transcription factor Her5 links patterning

to regional inhibition of neurogenesis at the midbrain-hindbrain boundary. *Development,* **130**, 1591-604.

Geling, A., Steiner, H., Willem, M., Bally-Cuif, L. & Haass, C. (2002). A gamma-secretase inhibitor blocks Notch signaling in vivo and causes a severe neurogenic phenotype in zebrafish. *EMBO Rep.,* **3**, 688-94.

Gerety, S. S., Breau, M. A., Sasai, N., Xu, Q., Briscoe, J. & Wilkinson, D. G. (2013). An inducible transgene expression system for zebrafish and chick. *Development,* **140**, 2235-43.

Gerety, S. S. & Wilkinson, D. G. (2011). Morpholino artifacts provide pitfalls and reveal a novel role for pro-apoptotic genes in hindbrain boundary development. *Dev. Biol.,* **350**, 279-89.

Ghysen, A. & Dambly-Chaudière, C. (2007). The lateral line microcosmos. *Genes Dev.,* **21**, 2118-30.

Giardine, B., Riemer, C., Hardison, R. C., Burhans, R., Elnitski, L., Shah, P., Zhang, Y., Blankenberg, D., Albert, I., Taylor, J., Miller, W., Kent, W. J. & Nekrutenko, A. (2005). Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.,* **15**, 1451-5.

Gleimer, M., Von Boehmer, H. & Kreslavsky, T. (2012). PLZF Controls the Expression of a Limited Number of Genes Essential for NKT Cell Function. *Frontiers in immunology,* **3**, 374-374.

Goecks, J., Nekrutenko, A. & Taylor, J. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.,* **11**, R86-R86.

Gonzalez-Quevedo, R., Lee, Y., Poss, K. D. & Wilkinson, D. G. (2010). Neuronal regulation of the spatial patterning of neurogenesis. *Dev. Cell,* **18**, 136-47.

Goodfellow, M., Phillips, N. E., Manning, C., Galla, T. & Papalopulu, N. (2014). microRNA input into a neural ultradian oscillator controls emergence and timing of alternative cell states. *Nat. Commun.,* **5**, 3399-3399.

Gradwohl, G., Fode, C. & Guillemot, F. (1996). Restricted expression of a novel murine atonal-related bHLH protein in undifferentiated neural precursors. *Dev. Biol.,* **180**, 227-41.

Graham, V., Khudyakov, J., Ellis, P. & Pevny, L. (2003). SOX2 functions to maintain neural progenitor identity. *Neuron,* **39**, 749-65.

Grandel, H., Kaslin, J., Ganz, J., Wenzel, I. & Brand, M. (2006). Neural stem cells and neurogenesis in the adult zebrafish brain: origin, proliferation dynamics, migration and cell fate. *Dev. Biol.,* **295**, 263-77.

Grandel, H. & Schulte-Merker, S. (1998). The development of the paired fins in the zebrafish (Danio rerio). *Mech. Dev.,* **79**, 99-120.

Grbavec, D. & Stifani, S. (1996). Molecular interaction between TLE1 and the carboxyl-terminal domain of HES-1 containing the WRPW motif. *Biochem. Biophys. Res. Commun.,* **223**, 701-5.

Grunwald, D. J., Kimmel, C. B., Westerfield, M., Walker, C. & Streisinger, G. (1988). A neural degeneration mutation that spares primary neurons in the zebrafish. *Dev. Biol.,* **126**, 115-28.

Guidez, F., Howell, L., Isalan, M., Cebrat, M., Alani, R. M., Ivins, S., Hormaeche, I., Mcconnell, M. J., Pierce, S., Cole, P. A., Licht, J. & Zelent, A. (2005). Histone acetyltransferase activity of p300 is required for transcriptional repression by the promyelocytic leukemia zinc finger protein. *Mol. Cell. Biol.,* **25**, 5552-66.

Guidez, F., Parks, S., Wong, H., Jovanovic, J. V., Mays, A., Gilkes, A. F., Mills, K. I., Guillemin, M.-C., Hobbs, R. M., Pandolfi, P. P., De Thé, H., Solomon, E. & Grimwade, D. (2007). RARalpha-PLZF overcomes PLZF-mediated repression of CRABPI, contributing to retinoid resistance in t(11;17) acute promyelocytic leukemia. *Proc. Natl. Acad. Sci. U. S. A.,* **104**, 18694-9.

Guilinger, J. P., Pattanayak, V., Reyon, D., Tsai, S. Q., Sander, J. D., Joung, J. K. & Liu, D. R. (2014). Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat. Methods,* **11**, 429-435.

Guillemot, F. (2007). Spatial and temporal specification of neural fates by transcription factor codes. *Development,* **134**, 3771-80.

Gundry, C. N., Vandersteen, J. G., Reed, G. H., Pryor, R. J., Chen, J. & Wittwer, C. T. (2003). Amplicon melting analysis with labeled primers: a closed-tube method for differentiating homozygotes and heterozygotes. *Clin. Chem.,* **49**, 396-406.

Haddon, C., Smithers, L., Schneider-Maunoury, S., Coche, T., Henrique, D. & Lewis, J. (1998). Multiple delta genes and lateral inhibition in zebrafish primary neurogenesis. *Development,* **125**, 359-70.

Hanneman, E., Trevarrow, B., Metcalfe, W. K., Kimmel, C. B. & Westerfield, M. (1988). Segmental pattern of development of the hindbrain and spinal cord of the zebrafish embryo. *Development,* **103**, 49-58.

Hardcastle, Z., Chalmers, A. D. & Papalopulu, N. (2000). FGF-8 stimulates neuronal differentiation through FGFR-4a and interferes with mesoderm induction in Xenopus embryos. *Curr. Biol.,* **10**, 1511-1514.

Hardcastle, Z. & Papalopulu, N. (2000). Distinct effects of XBF-1 in regulating the cell cycle inhibitor p27(XIC1) and imparting a neural fate. *Development,* **127**, 1303-14.

Harrison, M. R., Georgiou, A. S., Spaink, H. P. & Cunliffe, V. T. (2011). The epigenetic regulator Histone Deacetylase 1 promotes transcription of a core neurogenic programme in zebrafish embryos. *BMC Genomics,* **12**, 24-24.

Hatakeyama, J., Bessho, Y., Katoh, K., Ookawara, S., Fujioka, M., Guillemot, F. & Kageyama, R. (2004). Hes genes regulate size, shape and histogenesis of the nervous

system by control of the timing of neural stem cell differentiation. *Development,* **131**, 5539-50.

Hirata, H., Yoshiura, S., Ohtsuka, T., Bessho, Y., Harada, T., Yoshikawa, K. & Kageyama, R. (2002). Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop. *Science,* **298**, 840-3.

Hoatlin, M. E., Zhi, Y., Ball, H., Silvey, K., Melnick, A., Stone, S., Arai, S., Hawe, N., Owen, G., Zelent, A. & Licht, J. D. (1999). A novel BTB/POZ transcriptional repressor protein interacts with the Fanconi anemia group C protein and PLZF. *Blood,* **94**, 3737-47.

Hobbs, R. M., Fagoonee, S., Papa, A., Webster, K., Altruda, F., Nishinakamura, R., Chai, L. & Pandolfi, P. P. (2012). Functional Antagonism between Sall4 and Plzf Defines Germline Progenitors. *Cell Stem Cell,* **10**, 284-98.

Hobbs, R. M., Seandel, M., Falciatori, I., Rafii, S. & Pandolfi, P. P. (2010). Plzf Regulates Germline Progenitor Self-Renewal by Opposing mTORC1. *Cell,* **142**, 468-479.

Hong, S. H., David, G., Wong, C. W. & Dejean, A. (1997). SMRT corepressor interacts with PLZF and with the PML-retinoic acid receptor α (RARα) and PLZF-RARα oncoproteins associated with acute promyelocytic leukemia. *Proc. Natl. Acad. Sci. U. S. A.,* **94**, 9028-9033.

Hu, G., Goll, M. G. & Fisher, S. (2011). ΦC31 integrase mediates efficient cassette exchange in the zebrafish germline. *Dev. Dyn.,* **240**, 2101-7.

Huang, P., Xiao, A., Zhou, M., Zhu, Z., Lin, S. & Zhang, B. (2011). Heritable gene targeting in zebrafish using customized TALENs. *Nat. Biotechnol.,* **29**, 699-700.

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Kaini, P., Sander, J. D., Joung, J. K., Peterson, R. T. & Yeh, J.-R. J. (2013a). Heritable and precise zebrafish genome editing using a CRISPR-Cas system. *PLoS ONE,* **8**, e68708-e68708.

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J. R. J. & Joung, J. K. (2013b). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.*, 29-31.

Imayoshi, I., Isomura, A., Harima, Y., Kawaguchi, K., Kori, H., Miyachi, H., Fujiwara, T., Ishidate, F. & Kageyama, R. (2013). Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. *Science,* **342**, 1203-8.

Isshiki, T., Pearson, B., Holbrook, S. & Doe, C. Q. (2001). Drosophila Neuroblasts Sequentially Express Transcription Factors which Specify the Temporal Identity of Their Neuronal Progeny. *Cell,* **106**, 511-521.

Itoh, M., Kim, C.-H., Palardy, G., Oda, T., Jiang, Y. J., Maust, D., Yeo, S. Y., Lorick, K., Wright, G. J., Ariza-Mcnaughton, L. & Others (2003). Mind bomb is a ubiquitin ligase that is essential for efficient activation of Notch signaling by Delta. *Dev. Cell,* **4**, 67-82.

Itoh, M., Kudoh, T., Dedekian, M., Kim, C.-H. & Chitnis, A. B. (2002). A role for iro1 and iro7 in the establishment of an anteroposterior compartment of the ectoderm adjacent to the midbrain-hindbrain boundary. *Development,* **129**, 2317-27.

Ivins, S., Pemberton, K., Guidez, F., Howell, L., Krumlauf, R. & Zelent, A. (2003). Regulation of Hoxb2 by APL-associated PLZF protein. *Oncogene,* **22**, 3685-97.

Jacob, J., Maurange, C. & Gould, A. P. (2008). Temporal control of neuronal diversity: common regulatory principles in insects and vertebrates? *Development,* **135**, 3481-9.

Jan, Y. N. & Jan, L. Y. (1993). HLH proteins, fly neurogenesis, and vertebrate myogenesis. *Cell,* **75**, 827-30.

Jessell, T. M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nat. Rev. Genet.,* **1**, 20-9.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science,* **337**, 816-21.

Jones, C., St-Jean, S., Fréchette, I., Bergeron, D., Rivard, N. & Boudreau, F. (2013). Identification of a novel promyelocytic leukemia zinc-finger isoform required for colorectal cancer cell growth and survival. *Int. J. Cancer,* **133**, 58-66.

Kageyama, R., Ohtsuka, T., Hatakeyama, J. & Ohsawa, R. (2005). Roles of bHLH genes in neural stem cell differentiation. *Exp. Cell Res.,* **306**, 343-8.

Kang, S. I., Chang, W.-J., Cho, S.-G. & Kim, I. Y. (2003). Modification of promyelocytic leukemia zinc finger protein (PLZF) by SUMO-1 conjugation regulates its transcriptional repressor activity. *J. Biol. Chem.,* **278**, 51479-83.

Kang, S. I., Choi, H. W. & Kim, I. Y. (2008). Redox-mediated modification of PLZF by SUMO-1 and ubiquitin. *Biochem. Biophys. Res. Commun.,* **369**, 1209-14.

Kawakami, K., Takeda, H., Kawakami, N., Kobayashi, M., Matsuda, N. & Mishina, M. (2004). A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. *Dev. Cell,* **7**, 133-44.

Kidder, B. L., Hu, G. & Zhao, K. (2011). ChIP-Seq: technical considerations for obtaining high-quality data. *Nat. Immunol.,* **12**, 918-22.

Kiecker, C. & Lumsden, A. (2005). Compartments and their boundaries in vertebrate brain development. *Nat. Rev. Neurosci.,* **6**, 553-64.

Kiecker, C. & Lumsden, A. (2012). The role of organizers in patterning the nervous system. *Annu. Rev. Neurosci.,* **35**, 347-67.

Kim, C. H., Bae, Y. K., Yamanaka, Y., Yamashita, S., Shimizu, T., Fujii, R., Park, H. C., Yeo, S. Y., Huh, T. L., Hibi, M. & Hirano, T. (1997). Overexpression of neurogenin induces ectopic expression of HuC in zebrafish. *Neurosci. Lett.,* **239**, 113-6.

Kim, C. H., Ueshima, E., Muraoka, O., Tanaka, H., Yeo, S. Y., Huh, T. L. & Miki, N. (1996a). Zebrafish elav/HuC homologue as a very early neuronal marker. *Neurosci. Lett.,* **216**, 109-12.

Kim, E., Kim, S., Kim, D. H., Choi, B.-S., Choi, I.-Y. & Kim, J.-S. (2012). Precision genome engineering with programmable DNA-nicking enzymes. *Genome Res.,* **22**, 1327-33.

Kim, H. & Kim, J.-S. (2014). A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.,* **15**, 321-34.

Kim, H., Shin, J., Kim, S., Poling, J., Park, H.-C. & Appel, B. (2008). Notch-regulated oligodendrocyte specification from radial glia in the spinal cord of zebrafish embryos. *Dev. Dyn.,* **237**, 2081-9.

Kim, Y.-G., Cha, J. & Chandrasegaran, S. (1996b). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A.,* **93**, 1156-60.

Kimmel, C. B. (1993). Patterning the brain of the zebrafish embryo. *Annu. Rev. Neurosci.,* **16**, 707-32.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.,* **203**, 253-310.

Kimmel, C. B. & Westerfield, M. (1990). Primary neurons of the zebrafish. In*:* EDELMAN, G. M., GALL, W. E. & COWAN, M. W. (eds.) *Signal and Sense.* New York: Wiley-Liss, 561-588.

Kohwi, M. & Doe, C. Q. (2013). Temporal fate specification and neural progenitor competence during development. *Nat. Rev. Neurosci.,* **14**, 823-38.

Koken, M. H., Reid, A., Quignon, F., Chelbi-Alix, M. K., Davies, J. M., Kabarowski, J. H., Zhu, J., Dong, S., Chen, S., Chen, Z., Tan, C. C., Licht, J., Waxman, S., De Thé, H. & Zelent, A. (1997). Leukemia-associated retinoic acid receptor alpha fusion partners, PML and PLZF, heterodimerize and colocalize to nuclear bodies. *Proc. Natl. Acad. Sci. U. S. A.,* **94**, 10255-60.

Kondo, M. (2010). Lymphoid and myeloid lineage commitment in multipotent hematopoietic progenitors. *Immunol. Rev.,* **238**, 37-46.

Köprunner, M., Thisse, C., Thisse, B. & Raz, E. (2001). A zebrafish nanos-related gene is essential for the development of primordial germ cells. *Genes Dev.,* **15**, 2877-85.

Kotaja, N. & Sassone-Corsi, P. (2004). Plzf pushes stem cells. *Nat. Genet.,* **36**, 551-3.

Kovalovsky, D., Uche, O. U., Eladad, S., Hobbs, R. M., Yi, W., Alonzo, E., Chua, K., Eidson, M., Kim, H.-J., Im, J. S., Pandolfi, P. P. & Sant'angelo, D. B. (2008). The BTB-zinc finger transcriptional regulator PLZF controls the development of invariant natural killer T cell effector functions. *Nat. Immunol.,* **9**, 1055-64.

Koyano-Nakagawa, N., Kim, J., Anderson, D. & Kintner, C. (2000). Hes6 acts in a positive feedback loop with the neurogenins to promote neuronal differentiation. *Development,* **127**, 4203-16.

Kurita, R., Sagara, H., Aoki, Y., Link, B. A., Arai, K.-I. & Watanabe, S. (2003). Suppression of lens growth by αA-crystallin promoter-driven expression of diphtheria toxin results in disruption of retinal cell organization in zebrafish. *Dev. Biol.,* **255**, 113-127.

Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D. S., Parant, J. M., Yost, H. J., Kanki, J. P. & Chien, C.-B. (2007). The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.,* **236**, 3088-99.

Kyriakis, J. M. & Avruch, J. (1996). Sounding the Alarm: Protein Kinase Cascades Activated by Stress and Inflammation. *J. Biol. Chem.,* **271**, 24313-24316.

Labbaye, C., Quaranta, M. T., Pagliuca, A., Militi, S., Licht, J. D., Testa, U. & Peschle, C. (2002). PLZF induces megakaryocytic development, activates Tpo receptor expression and interacts with GATA1 protein. *Oncogene,* **21**, 6669-79.

Lam, C. S., März, M. & Strähle, U. (2009). Gfap and Nestin Reporter Lines Reveal Characteristics of Neural Progenitors in the Adult Zebrafish Brain. *Dev. Dyn.,* **238**, 475-86.

Lamason, R. L., Mohideen, M.-a. P. K., Mest, J. R., Wong, A. C., Norton, H. L., Aros, M. C., Jurynec, M. J., Mao, X., Humphreville, V. R., Humbert, J. E., Sinha, S., Moore, J. L., Jagadeeswaran, P., Zhao, W., Ning, G., Makalowska, I., Mckeigue, P. M., O'donnell, D., Kittles, R., Parra, E. J., Mangini, N. J., Grunwald, D. J., Shriver, M. D., Canfield, V. A. & Cheng, K. C. (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science,* **310**, 1782-6.

Lamb, B. M., Mercer, A. C. & Barbas, C. F. (2013). Directed evolution of the TALE N-terminal domain for recognition of all 5' bases. *Nucleic Acids Res.,* **41**, 9779-85.

Lange, C., Huttner, W. B. & Calegari, F. (2009). Cdk4/cyclinD1 overexpression in neural stem cells shortens G1, delays neurogenesis, and promotes the generation and expansion of basal progenitors. *Cell Stem Cell,* **5**, 320-31.

Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.,* **10**, R25-R25.

Lauter, G., Söll, I. & Hauptmann, G. (2011). Two-color fluorescent in situ hybridization in the embryonic zebrafish brain using differential detection systems. *BMC Dev. Biol.,* **11**, 43-43.

Law, S. H. W. & Sargent, T. D. (2013). Maternal pak4 expression is required for primitive myelopoiesis in zebrafish. *Mech. Dev.,* **130**, 181-94.

Law, S. H. W. & Sargent, T. D. (2014). The Serine-Threonine Protein Kinase PAK4 Is Dispensable in Zebrafish: Identification of a Morpholino-Generated Pseudophenotype. *PLoS ONE,* **9**, e100268-e100268.

Lee, J., Hollenberg, S., Snider, L., Turner, D., Lipnick, N. & Weintraub, H. (1995). Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. *Science,* **268**, 836-844.

Li, J. Y., English, M. A., Ball, H. J., Yeyati, P. L., Waxman, S. & Licht, J. D. (1997). Sequence-specific DNA binding and transcriptional regulation by the promyelocytic leukemia zinc finger protein. *J. Biol. Chem.,* **272**, 22447-55.

Li, X., Peng, H., Schultz, D. C., Lopez-Guisa, J. M., Rauscher, F. J. & Marmorstein, R. (1999). Structure-function studies of the BTB/POZ transcriptional repression domain from the promyelocytic leukemia zinc finger oncoprotein. *Cancer Res.,* **59**, 5275-82.

Licht, J. D., Chomienne, C., Goy, A., Chen, A., Scott, A. A., Head, D. R., Michaux, J. L., Wu, Y., Deblasio, A. & Miller, W. H. (1995). Clinical and molecular characterization of a rare syndrome of acute promyelocytic leukemia associated with translocation (11;17). *Blood,* **85**, 1083-94.

Liew, M., Pryor, R., Palais, R., Meadows, C., Erali, M., Lyon, E. & Wittwer, C. (2004). Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. *Clin. Chem.,* **50**, 1156-64.

Lin, D.-Y., Huang, C.-C., Hsieh, Y.-T., Lin, H.-C., Pao, P.-C., Tsou, J.-H., Lai, C.-Y., Hung, L.-Y., Wang, J.-M., Chang, W.-C. & Lee, Y.-C. (2013). Analysis of the interaction between Zinc finger protein 179 (Znf179) and promyelocytic leukemia zinc finger (Plzf). *J. Biomed. Sci.,* **20**, 98-98.

Lin, R. J., Nagy, L., Inoue, S., Shao, W., Miller, W. H. & Evans, R. M. (1998). Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature,* **391**, 811-4.

Lin, R. J., Sternsdorf, T., Tini, M. & Evans, R. M. (2001). Transcriptional regulation in acute promyelocytic leukemia. *Oncogene,* **20**, 7204-15.

Lin, S. & Lee, T. (2012). Generating neuronal diversity in the Drosophila central nervous system. *Dev. Dyn.,* **241**, 57-68.

Lin, Y., Fine, E. J., Zheng, Z., Antico, C. J., Voit, R. A., Porteus, M. H., Cradick, T. J. & Bao, G. (2014). SAPTA: a new design tool for improving TALE nuclease activity. *Nucleic Acids Res.,* **42**, e47-e47.

Liska, F., Snajdr, P., Sedová, L., Seda, O., Chylíková, B., Slámová, P., Krejcí, E., Sedmera, D., Grim, M., Krenová, D. & Kren, V. (2009). Deletion of a conserved noncoding sequence in Plzf intron leads to Plzf down-regulation in limb bud and polydactyly in the rat. *Dev. Dyn.,* **238**, 673-84.

Liu, E. T., Pott, S. & Huss, M. (2010). Q & A: ChIP-seq technologies and the study of gene regulation. *BMC Biol.,* **8**, 56-56.

Liu, K. J. & Harland, R. M. (2005). Inhibition of neurogenesis by SRp38, a neuroD-regulated RNA-binding protein. *Development,* **132**, 1511-23.

Lumsden, A. & Krumlauf, R. (1996). Patterning the vertebrate neuraxis. *Science,* **274**, 1109-15.

Lyden, D., Young, A. Z., Zagzag, D., Yan, W., Gerald, W., O'reilly, R., Bader, B. L., Hynes, R. O., Zhuang, Y., Manova, K. & Benezra, R. (1999). Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts. *Nature,* **401**, 670-7.

Lyons, D. A., Guy, A. T. & Clarke, J. D. W. (2003). Monitoring neural progenitor fate through multiple rounds of division in an intact vertebrate brain. *Development,* **130**, 3427-36.

Ma, Q., Kintner, C. & Anderson, D. J. (1996). Identification of neurogenin, a vertebrate neuronal determination gene. *Cell,* **87**, 43-52.

Maden, M. (2007). Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nat. Rev. Neurosci.,* **8**, 755-65.

Mak, A. N.-S., Bradley, P., Cernadas, R. A., Bogdanove, A. J. & Stoddard, B. L. (2012). The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science,* **335**, 716-9.

Malagelada, C., López-Toledano, M. A., Willett, R. T., Jin, Z. H., Shelanski, M. L. & Greene, L. A. (2011). RTP801/REDD1 Regulates the Timing of Cortical Neurogenesis and Neuron Migration. *J. Neurosci.,* **31**, 3186-96.

Malatesta, P., Appolloni, I. & Calzolari, F. (2008). Radial glia and neural stem cells. *Cell Tissue Res.,* **331**, 165-78.

Marcus, R. C. & Easter, S. S. (1995). Expression of glial fibrillary acidic protein and its relation to tract formation in embryonic zebrafish (Danio rerio). *J. Comp. Neurol.,* **359**, 365-81.

Maresca, M., Lin, V. G., Guo, N. & Yang, Y. (2013). Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. *Genome Res.,* **23**, 539-46.

Martin, P. J., Delmotte, M.-H., Formstecher, P. & Lefebvre, P. (2003). PLZF is a negative regulator of retinoic acid receptor transcriptional activity. *Nuclear receptor,* **1**, 6-6.

Massari, M. E. & Murre, C. (2000). Helix-Loop-Helix Proteins: Regulators of Transcription in Eucaryotic Organisms. *Mol. Cell. Biol.,* **20**, 429-440.

Mathew, R., Seiler, M. P., Scanlon, S. T., Mao, A.-P., Constantinides, M. G., Bertozzi-Villa, C., Singer, J. D. & Bendelac, A. (2012). BTB-ZF factors recruit the E3 ligase cullin 3 to regulate lymphoid effector programs. *Nature*, 2-7.

Mcconnell, M. J., Chevallier, N., Berkofsky-Fessler, W., Giltnane, J. M., Malani, R. B., Staudt, L. M. & Licht, J. D. (2003). Growth suppression by acute promyelocytic leukemia-associated protein PLZF is mediated by repression of c-myc expression. *Mol. Cell. Biol.,* **23**, 9375-88.

Mcconnell, M. J. & Licht, J. D. (2007). The PLZF Gene of t(11;17)-Associated APL. In*:* PANDOLFI, P. & VOGT, P. (eds.) *Acute Promyelocytic Leukemia.* Springer Berlin Heidelberg, 31-48.

Melnick, A., Ahmad, K. F., Arai, S., Polinger, A., Ball, H., Borden, K. L., Carlile, G. W., Prive, G. G. & Licht, J. D. (2000a). In-depth mutational analysis of the promyelocytic leukemia zinc finger BTB/POZ domain reveals motifs and residues required for biological and transcriptional functions. *Mol. Cell. Biol.,* **20**, 6550-67.

Melnick, A., Carlile, G., Ahmad, K. F., Kiang, C. L., Corcoran, C., Bardwell, V., Prive, G. G. & Licht, J. D. (2002). Critical residues within the BTB domain of PLZF and Bcl-6 modulate interaction with corepressors. *Mol. Cell. Biol.,* **22**, 1804-1804.

Melnick, A., Carlile, G. W., Mcconnell, M. J., Polinger, A., Hiebert, S. W. & Licht, J. D. (2000b). AML-1/ETO fusion protein is a dominant negative inhibitor of transcriptional repression by the promyelocytic leukemia zinc finger protein. *Blood,* **96**, 3939-47.

Melnick, A. M., Westendorf, J. J., Polinger, A., Carlile, G. W., Arai, S., Ball, H. J., Lutterbach, B., Hiebert, S. W. & Licht, J. D. (2000c). The ETO protein disrupted in t(8;21)-associated acute myeloid leukemia is a corepressor for the promyelocytic leukemia zinc finger protein. *Mol. Cell. Biol.,* **20**, 2075-86.

Mendelson, B. (1986). Development of reticulospinal neurons of the zebrafish. II. Early axonal outgrowth and cell body position. *J. Comp. Neurol.,* **251**, 172-84.

Metzger, D. & Chambon, P. (2001). Site- and time-specific gene targeting in the mouse. *Methods,* **24**, 71-80.

Michaelidis, T. M. & Lie, D. C. (2008). Wnt signaling and neural stem cells: caught in the Wnt web. *Cell Tissue Res.,* **331**, 193-210.

Miller, F. D. & Gauthier, A. S. (2007). Timing is everything: making neurons versus glia in the developing cortex. *Neuron,* **54**, 357-69.

Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y.-L., Rupniewski, I., Beausejour, C. M., Waite, A. J., Wang, N. S., Kim, K. A., Gregory, P. D., Pabo, C. O. & Rebar, E. J. (2007). An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.,* **25**, 778-85.

Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., Meng, X., Paschon, D. E., Leung, E., Hinkley, S. J., Dulay, G. P., Hua, K. L., Ankoudinova, I., Cost, G. J., Urnov, F. D., Zhang, H. S., Holmes, M. C., Zhang, L., Gregory, P. D. & Rebar, E. J. (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.,* **29**, 143-8.

Ming, G.-L. & Song, H. (2011). Adult neurogenesis in the mammalian brain: significant answers and significant questions. *Neuron,* **70**, 687-702.

Misra, K., Mishra, K., Gui, H. & Matise, M. P. (2008). Prox1 regulates a transitory state for interneuron neurogenesis in the spinal cord. *Dev. Dyn.,* **237**, 393-402.

Mitchelmore, C., Kjaerulff, K. M., Pedersen, H. C., Nielsen, J. V., Rasmussen, T. E., Fisker, M. F., Finsen, B., Pedersen, K. M. & Jensen, N. A. (2002). Characterization of two novel nuclear BTB/POZ domain zinc finger isoforms. Association with differentiation of hippocampal neurons, cerebellar granule cells, and macroglia. *J. Biol. Chem.,* **277**, 7598-609.

Mizuguchi, R., Sugimori, M., Takebayashi, H., Kosako, H., Nagao, M., Yoshida, S., Nabeshima, Y., Shimamura, K. & Nakafuku, M. (2001). Combinatorial roles of olig2 and neurogenin2 in the coordinated induction of pan-neuronal and subtype-specific properties of motoneurons. *Neuron,* **31**, 757-71.

Moloney, D. J., Panin, V. M., Johnston, S. H., Chen, J., Shao, L., Wilson, R., Wang, Y., Stanley, P., Irvine, K. D., Haltiwanger, R. S. & Vogt, T. F. (2000). Fringe is a glycosyltransferase that modifies Notch. *Nature,* **406**, 369-75.

Montgomery, R. L., Hsieh, J., Barbosa, A. C., Richardson, J. A. & Olson, E. N. (2009). Histone deacetylases 1 and 2 control the progression of neural precursors to neurons during brain development. *Proc. Natl. Acad. Sci. U. S. A.,* **106**, 7876-81.

Morris, D. R. & Geballe, A. P. (2000). Upstream Open Reading Frames as Regulators of mRNA Translation. *Mol. Cell. Biol.,* **20**, 8635-8642.

Moscou, M. J. & Bogdanove, A. J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science,* **326**, 1501-1501.

Murray, A. W. (2004). Recycling the cell cycle: cyclins revisited. *Cell,* **116**, 221-34.

Murray-Zmijewski, F., Lane, D. P. & Bourdon, J. C. (2006). P53/P63/P73 Isoforms: an Orchestra of Isoforms To Harmonise Cell Differentiation and Response To Stress. *Cell Death Differ.,* **13**, 962-72.

Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T. & Cathomen, T. (2011). A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.,* **39**, 9283-93.

Nanba, D., Mammoto, A., Hashimoto, K. & Higashiyama, S. (2003). Proteolytic release of the carboxy-terminal fragment of proHB-EGF causes nuclear export of PLZF. *J. Cell Biol.,* **163**, 489-502.

Nikolaou, N., Watanabe-Asaka, T., Gerety, S., Distel, M., Köster, R. W. & Wilkinson, D. G. (2009). Lunatic fringe promotes the lateral inhibition of neurogenesis. *Development,* **136**, 2523-33.

Noctor, S. C., Flint, A. C., Weissman, T. A., Wong, W. S., Clinton, B. K. & Kriegstein, A. R. (2002). Dividing precursor cells of the embryonic cortical ventricular zone have morphological and molecular characteristics of radial glia. *J. Neurosci.,* **22**, 3161-73.

Noctor, S. C., Martínez-Cerdeño, V., Ivic, L. & Kriegstein, A. R. (2004). Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat. Neurosci.,* **7**, 136-44.

Odom, D. T., Dowell, R. D., Jacobsen, E. S., Gordon, W., Danford, T. W., Macisaac, K. D., Rolfe, P. A., Conboy, C. M., Gifford, D. K. & Fraenkel, E. (2007). Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat. Genet.,* **39**, 730-2.

Orlando, S. J., Santiago, Y., Dekelver, R. C., Freyvert, Y., Boydston, E. A., Moehle, E. A., Choi, V. M., Gopalan, S. M., Lou, J. F., Li, J., Miller, J. C., Holmes, M. C., Gregory, P. D., Urnov, F. D. & Cost, G. J. (2010). Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res.,* **38**, e152-e152.

Ozaki, Y., Saito, K., Shinya, M., Kawasaki, T. & Sakai, N. (2011). Evaluation of Sycp3, Plzf and Cyclin B3 expression and suitability as spermatogonia and spermatocyte markers in zebrafish. *Gene expression patterns : GEP,* **11**, 309-15.

Pagliuca, A., Gallo, P., De Luca, P. & Lania, L. (2000). Class A helix-loop-helix proteins are positive regulators of several cyclin-dependent kinase inhibitors' promoter activity and negatively affect cell growth. *Cancer Res.,* **60**, 1376-82.

Papan, C. & Campos-Ortega, J. (1994). On the formation of the neural keel and neural tube in the zebrafish Danio (Brachydanio) rerio. *Roux's Archives of Developmental Biology,* **203**, 178-186.

Parrado, A., Robledo, M., Moya-Quiles, M. R., Marín, L. A., Chomienne, C., Padua, R. A. & Alvarez-López, M. R. (2004). The promyelocytic leukemia zinc finger protein down-regulates apoptosis and expression of the proapoptotic BID protein in lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.,* **101**, 1898-903.

Pascale, A., Amadio, M. & Quattrone, A. (2008). Defining a neuron: neuronal ELAV proteins. *Cellular and molecular life sciences : CMLS,* **65**, 128-40.

Pasini, A. & Wilkinson, D. G. (2002). Stabilizing the regionalisation of the developing vertebrate central nervous system. *Bioessays,* **24**, 427-38.

Pellegrini, E., Mouriec, K., Anglade, I., Menuet, A., Le Page, Y., Gueguen, M.-M., Marmignon, M.-H., Brion, F., Pakdel, F. & Kah, O. (2007). Identification of aromatase-positive radial glial cells as progenitor cells in the ventricular layer of the forebrain in zebrafish. *J. Comp. Neurol.,* **501**, 150-67.

Perrier, A. L., Tabar, V., Barberi, T., Rubio, M. E., Bruses, J., Topf, N., Harrison, N. L. & Studer, L. (2004). Derivation of midbrain dopamine neurons from human embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.,* **101**, 12543-8.

Perron, M., Kanekar, S., Vetter, M. L. & Harris, W. A. (1998). The genetic sequence of retinal development in the ciliary margin of the Xenopus eye. *Dev. Biol.,* **199**, 185-200.

Perron, M., Opdecamp, K., Butler, K., Harris, W. A. & Bellefroid, E. J. (1999). X-ngnr-1 and Xath3 promote ectopic expression of sensory neuron markers in the neurula ectoderm and have distinct inducing properties in the retina. *Proc. Natl. Acad. Sci. U. S. A.,* **96**, 14996-5001.

Petrie, K., Guidez, F., Zhu, J., Howell, L., Owen, G., Chew, Y. P., Parks, S., Waxman, S., Licht, J., Mittnacht, S. & Zelent, A. (2008). Retinoblastoma protein and the leukemia-associated PLZF transcription factor interact to repress target gene promoters. *Oncogene,* **27**, 5260-6.

Pevny, L. & Placzek, M. (2005). SOX genes and neural progenitor identity. *Curr. Opin. Neurobiol.,* **15**, 7-13.

Pilz, G.-A., Shitamukai, A., Reillo, I., Pacary, E., Schwausch, J., Stahl, R., Ninkovic, J., Snippert, H. J., Clevers, H., Godinho, L., Guillemot, F., Borrell, V., Matsuzaki, F. & Götz, M. (2013). Amplification of progenitors in the mammalian telencephalon includes a new radial glial cell type. *Nat. Commun.,* **4**, 2125-2125.

Poliakov, A., Cotrina, M. L., Pasini, A. & Wilkinson, D. G. (2008). Regulation of EphB2 activation and cell repulsion by feedback control of the MAPK pathway. *J. Cell Biol.,* **183**, 933-47.

Postlethwait, J. H., Yan, Y. L., Gates, M. A., Horne, S., Amores, A., Brownlie, A., Donovan, A., Egan, E. S., Force, A., Gong, Z., Goutel, C., Fritz, A., Kelsh, R., Knapik, E., Liao, E., Paw, B., Ransom, D., Singer, A., Thomson, M., Abduljabbar, T. S., Yelick, P., Beier, D., Joly, J. S., Larhammar, D., Rosa, F., Westerfield, M., Zon, L. I., Johnson, S. L. & Talbot, W. S. (1998). Vertebrate genome evolution and the zebrafish gene map. *Nat. Genet.,* **18**, 345-9.

Poulain, M. & Ober, E. A. (2011). Interplay between Wnt2 and Wnt2bb controls multiple steps of early foregut-derived organ development. *Development,* **138**, 3557-68.

Privé, G. G., Melnick, A., Ahmad, K. F. & Licht, J. D. (2005). The BTB domain zinc finger proteins. In*:* SHIRO, I. & KULDELL, N. (eds.) *Zinc Finger Proteins: From Atomic Contact to Cellular Function.* Georgetown: Landes Biosciences, 134-150.

Qian, X., Shen, Q., Goderie, S. K., He, W., Capela, A., Davis, A. A. & Temple, S. (2000). Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. *Neuron,* **28**, 69-80.

Qiu, X., Lim, C.-H., Ho, S. H.-K., Lee, K.-H. & Jiang, Y.-J. (2009). Temporal Notch activation through Notch1a and Notch3 is required for maintaining zebrafish rhombomere boundaries. *Dev. Genes Evol.,* **219**, 339-51.

Rakic, P. (1972). Mode of cell migration to the superficial layers of fetal monkey neocortex. *J. Comp. Neurol.,* **145**, 61-83.

Rakic, P. (2003). Developmental and evolutionary adaptations of cortical radial glia. *Cereb. Cortex,* **13**, 541-9.

Ramirez, C. L., Certo, M. T., Mussolino, C., Goodwin, M. J., Cradick, T. J., Mccaffrey, A. P., Cathomen, T., Scharenberg, A. M. & Joung, J. K. (2012). Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res.,* **40**, 5560-8.

Ramirez, C. L., Foley, J. E., Wright, D. A., Müller-Lerch, F., Rahman, S. H., Cornu, T. I., Winfrey, R. J., Sander, J. D., Fu, F., Townsend, J. A., Cathomen, T., Voytas, D. F. & Joung, J. K. (2008). Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat. Methods,* **5**, 374-5.

Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y. & Zhang, F. (2013). Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell,* **154**, 1380-9.

Reid, A., Gould, A., Brand, N., Cook, M., Strutt, P., Li, J., Licht, J., Waxman, S., Krumlauf, R. & Zelent, A. (1995). Leukemia translocation gene, PLZF, is expressed with a speckled nuclear pattern in early hematopoietic progenitors. *Blood,* **86**, 4544-52.

Reyes, R., Haendel, M., Grant, D., Melancon, E. & Eisen, J. S. (2004). Slow degeneration of zebrafish Rohon-Beard neurons during programmed cell death. *Dev. Dyn.,* **229**, 30-41.

Reyon, D., Tsai, S. Q., Khayter, C., Foden, J. A., Sander, J. D. & Joung, J. K. (2012). FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.,* **30**, 460-5.

Rho, S. B., Chung, B. M. & Lee, J. H. (2007). TIMP-1 regulates cell proliferation by interacting with the ninth zinc finger domain of PLZF. *J. Cell. Biochem.,* **101**, 57-67.

Rice, K. L., Hormaeche, I., Doulatov, S., Flatow, J. M., Grimwade, D., Mills, K. I., Leiva, M., Ablain, J., Ambardekar, C., Mcconnell, M. J., Dick, J. E. & Licht, J. D. (2009). Comprehensive genomic screens identify a role for PLZF-RARalpha as a positive regulator of cell proliferation via direct regulation of c-MYC. *Blood,* **114**, 5499-511.

Riu, E., Chen, Z.-Y., Xu, H., He, C.-Y. & Kay, M. A. (2007). Histone modifications are associated with the persistence or silencing of vector-mediated transgene expression in vivo. *Mol. Ther.,* **15**, 1348-55.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. & Mesirov, J. P. (2011). Integrative genomics viewer. *Nat. Biotechnol.,* **29**, 24-6.

Robu, M. E., Larson, J. D., Nasevicius, A., Beiraghi, S., Brenner, C., Farber, S. A. & Ekker, S. C. (2007). P53 Activation By Knockdown Technologies. *PLoS Genet.,* **3**, e78-e78.

Roegiers, F. & Jan, Y. N. (2004). Asymmetric cell division. *Curr. Opin. Cell Biol.,* **16**, 195-205.

Römer, P., Hahn, S., Jordan, T., Strauss, T., Bonas, U. & Lahaye, T. (2007). Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science,* **318**, 645-8.

Rothenaigner, I., Krecsmarik, M., Hayes, J. A., Bahn, B., Lepier, A., Fortin, G., Götz, M., Jagasia, R. & Bally-Cuif, L. (2011). Clonal analysis by distinct viral vectors identifies bona fide neural stem cells in the adult zebrafish telencephalon and characterizes their division properties and fate. *Development,* **138**, 1459-69.

Rouet, P., Smih, F. & Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol.,* **14**, 8096-106.

Roztocil, T., Matter-Sadzinski, L., Alliod, C., Ballivet, M. & Matter, J. M. (1997). NeuroM, a neural helix-loop-helix transcription factor, defines a new transition stage in neurogenesis. *Development,* **124**, 3263-72.

Ruthardt, M., Orleth, A., Tomassoni, L., Puccetti, E., Riganelli, D., Alcalay, M., Mannucci, R., Nicoletti, I., Grignani, F., Fagioli, M. & Pelicci, P. G. (1998). The acute promyelocytic leukaemia specific PML and PLZF proteins localize to adjacent and functionally distinct nuclear bodies. *Oncogene,* **16**, 1945-53.

Ruzinova, M. B. & Benezra, R. (2003). Id proteins in development, cell cycle and cancer. *Trends Cell Biol.,* **13**, 410-418.

Sandberg, M., Källström, M. & Muhr, J. (2005). Sox21 promotes the progression of vertebrate neurogenesis. *Nat. Neurosci.,* **8**, 995-1001.

Sander, J. D., Cade, L., Khayter, C., Reyon, D., Peterson, R. T., Joung, J. K. & Yeh, J.-R. J. (2011). Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat. Biotechnol.,* **29**, 697-698.

Sanders, K. L., Catto, L. E., Bellamy, S. R. W. & Halford, S. E. (2009). Targeting individual subunits of the FokI restriction endonuclease to specific DNA strands. *Nucleic Acids Res.,* **37**, 2105-15.

Sasai, Y. (1998). Identifying the Missing LinksGenes that Connect Neural Induction and Primary Neurogenesis in Vertebrate Embryos. *Neuron,* **21**, 455-458.

Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R. & Nakanishi, S. (1992). Two mammalian helix-loop-helix factors structurally related to Drosophila hairy and Enhancer of split. *Genes Dev.,* **6**, 2620-2634.

Savage, A. K., Constantinides, M. G., Han, J., Picard, D., Martin, E., Li, B., Lantz, O. & Bendelac, A. (2008). The transcription factor PLZF directs the effector program of the NKT cell lineage. *Immunity,* **29**, 391-403.

Schmidt, D., Wilson, M. D., Ballester, B., Schwalie, P. C., Brown, G. D., Marshall, A., Kutter, C., Watt, S., Martinez-Jimenez, C. P., Mackay, S., Talianidis, I., Flicek, P. & Odom, D. T. (2010). Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science,* **328**, 1036-40.

Schmidt, D., Wilson, M. D., Spyrou, C., Brown, G. D., Hadfield, J. & Odom, D. T. (2009). ChIP-seq: using high-throughput sequencing to discover protein-DNA interactions. *Methods,* **48**, 240-8.

Schulte-Merker, S. & Stainier, D. Y. R. (2014). Out with the old, in with the new: reassessing morpholino knockdowns in light of genome editing technology. *Development,* **141**, 3103-3104.

Schwamborn, J. C., Berezikov, E. & Knoblich, J. A. (2009). The TRIM-NHL protein TRIM32 activates microRNAs and prevents self-renewal in mouse neural progenitors. *Cell,* **136**, 913-25.

Segal, D. J. & Meckler, J. F. (2013). Genome engineering at the dawn of the golden age. *Annu. Rev. Genomics Hum. Genet.,* **14**, 135-58.

Seidel, K., Kirsch, S., Lucht, K., Zaade, D., Reinemund, J., Schmitz, J., Klare, S., Li, Y., Schefe, J. H., Schmerbach, K., Goldin-Lang, P., Zollmann, F. S., Thöne-Reineke, C., Unger, T. & Funke-Kaiser, H. (2010). The Promyelocytic Leukemia Zinc Finger (PLZF) Protein Exerts Neuroprotective Effects in Neuronal Cells and is Dysregulated in Experimental Stroke. *Brain Pathol.,* **21**, 31-43.

Seiler, M. P., Mathew, R., Liszewski, M. K., Spooner, C., Barr, K., Meng, F., Singh, H. & Bendelac, A. (2012). Elevated and sustained expression of the transcription factors Egr1 and Egr2 controls NKT lineage differentiation in response to TCR signaling. *Nat. Immunol.,* **13**, 264-71.

Seo, S., Lim, J.-W., Yellajoshyula, D., Chang, L.-W. & Kroll, K. L. (2007). Neurogenin and NeuroD direct transcriptional targets and their regulatory enhancers. *EMBO J.,* **26**, 5093-108.

Shaknovich, R., Yeyati, P. L., Ivins, S., Melnick, A., Lempert, C., Waxman, S., Zelent, A. & Licht, J. D. (1998). The promyelocytic leukemia zinc finger protein affects myeloid cell growth, differentiation, and apoptosis. *Mol. Cell. Biol.,* **18**, 5533-45.

Sham, M. H., Vesque, C., Nonchev, S., Marshall, H., Frain, M., Gupta, R. D., Whiting, J., Wilkinson, D., Charnay, P. & Krumlauf, R. (1993). The zinc finger gene Krox20 regulates HoxB2 (Hox2.8) during hindbrain segmentation. *Cell,* **72**, 183-96.

Shimojo, H., Ohtsuka, T. & Kageyama, R. (2008). Oscillations in notch signaling regulate maintenance of neural progenitors. *Neuron,* **58**, 52-64.

Sitterlin, D., Tiollais, P. & Transy, C. (1997). The RAR alpha-PLZF chimera associated with Acute Promyelocytic Leukemia has retained a sequence-specific DNA-binding domain. *Oncogene,* **14**, 1067-74.

Sobieszczuk, D. F., Poliakov, A., Xu, Q. & Wilkinson, D. G. (2010). A feedback loop mediated by degradation of an inhibitor is required to initiate neuronal differentiation. *Genes Dev.,* **24**, 206-18.

Sprague, J., Bayraktaroglu, L., Clements, D., Conlin, T., Fashena, D., Frazer, K., Haendel, M., Howe, D. G., Mani, P., Ramachandran, S., Schaper, K., Segerdell, E., Song, P., Sprunger, B., Taylor, S., Van Slyke, C. E. & Westerfield, M. (2006). The Zebrafish Information Network: the zebrafish model organism database. *Nucleic Acids Res.,* **34**, D581-5.

Stigloher, C., Chapouton, P., Adolf, B. & Bally-Cuif, L. (2008). Identification of neural progenitor pools by E(Spl) factors in the embryonic and adult brain. *Brain Res. Bull.,* **75**, 266-73.

Suliman, B. A., Xu, D. & Williams, B. R. G. (2012). The promyelocytic leukemia zinc finger protein: two decades of molecular oncology. *Frontiers in oncology,* **2**, 74-74.

Summerton, J. & Weller, D. (1997). Morpholino antisense oligomers: design, preparation, and properties. *Antisense Nucleic Acid Drug Dev.,* **7**, 187-95.

Sun, Y., Nadal-Vicens, M., Misono, S., Lin, M. Z., Zubiaga, A., Hua, X., Fan, G. & Greenberg, M. E. (2001). Neurogenin promotes neurogenesis and inhibits glial differentiation by independent mechanisms. *Cell,* **104**, 365-76.

Suster, M. L., Sumiyama, K. & Kawakami, K. (2009). Transposon-mediated BAC transgenesis in zebrafish and mice. *BMC Genomics,* **10**, 477-477.

Szczepek, M., Brondani, V., Büchel, J., Serrano, L., Segal, D. J. & Cathomen, T. (2007). Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.,* **25**, 786-93.

Tailor, J., Kittappa, R., Leto, K., Gates, M., Borel, M., Paulsen, O., Spitzer, S., Karadottir, R. T., Rossi, F., Falk, A. & Smith, A. (2013). Stem cells expanded from the human embryonic hindbrain stably retain regional specification and high neurogenic potency. *J. Neurosci.,* **33**, 12407-22.

Takahashi, M. & Osumi, N. (2011). Pax6 regulates boundary-cell specification in the rat hindbrain. *Mech. Dev.*

Takahashi, S. & Licht, J. D. (2002). The human promyelocytic leukemia zinc finger gene is regulated by the Evi-1 oncoprotein and a novel guanine-rich site binding protein. *Leukemia,* **16**, 1755-62.

Takahashi, T., Nowakowski, R. S. & Caviness, V. S. (1993). Cell cycle parameters and patterns of nuclear movement in the neocortical proliferative zone of the fetal mouse. *J. Neurosci.,* **13**, 820-33.

Tanabe, Y. & Jessell, T. M. (1996). Diversity and Pattern in the Developing Spinal Cord. *Science,* **274**, 1115-1123.

Tanigaki, K., Nogaki, F., Takahashi, J., Tashiro, K., Kurooka, H. & Honjo, T. (2001). Notch1 and Notch3 instructively restrict bFGF-responsive multipotent neural progenitor cells to an astroglial fate. *Neuron,* **29**, 45-55.

Tawk, M., Araya, C., Lyons, D. A., Reugels, A. M., Girdler, G. C., Bayley, P. R., Hyde, D. R., Tada, M. & Clarke, J. D. W. (2007). A mirror-symmetric cell division that orchestrates neuroepithelial morphogenesis. *Nature,* **446**, 797-800.

Terriente, J., Gerety, S. S., Watanabe-Asaka, T., Gonzalez-Quevedo, R. & Wilkinson, D. G. (2012). Signalling from hindbrain boundaries regulates neuronal clustering that patterns neurogenesis. *Development,* **139**, 2978-87.

Torii, M., Matsuzaki, F., Osumi, N., Kaibuchi, K., Nakamura, S., Casarosa, S., Guillemot, F. & Nakafuku, M. (1999). Transcription factors Mash-1 and Prox-1 delineate early steps in differentiation of neural stem cells in the developing central nervous system. *Development,* **126**, 443-56.

Trevarrow, B., Marks, D. L. & Kimmel, C. B. (1990). Organization of Hindbrain Segments in the Zebrafish Embryo. *Neuron,* **4**, 669-679.

Trinh, L. A., Hochgreb, T., Graham, M., Wu, D., Ruf-Zamojski, F., Jayasena, C. S., Saxena, A., Hawk, R., Gonzalez-Serricchio, A., Dixson, A., Chow, E., Gonzales, C., Leung, H.-Y., Solomon, I., Bronner-Fraser, M., Megason, S. G. & Fraser, S. E. (2011). A versatile gene trap to visualize and interrogate the function of the vertebrate proteome. *Genes Dev.,* **25**, 2306-20.

Tsuzuki, S. & Enver, T. (2002). Interactions of GATA-2 with the promyelocytic leukemia zinc finger (PLZF) protein, its homologue FAZF, and the t(11;17)-generated PLZF-retinoic acid receptor alpha oncoprotein. *Blood,* **99**, 3404-3410.

Vasquez, K. M., Marburger, K., Intody, Z. & Wilson, J. H. (2001). Manipulating the mammalian genome by homologous recombination. *Proc. Natl. Acad. Sci. U. S. A.,* **98**, 8403-10.

Vilas-Boas, F. & Henrique, D. (2010). HES6-1 and HES6-2 function through different mechanisms during neuronal differentiation. *PLoS ONE,* **5**, e15459-e15459.

Villa, R., Pasini, D., Gutierrez, A., Morey, L., Occhionorelli, M., Viré, E., Nomdedeu, J. F., Jenuwein, T., Pelicci, P. G., Minucci, S., Fuks, F., Helin, K. & Di Croce, L. (2007). Role of the polycomb repressive complex 2 in acute promyelocytic leukemia. *Cancer Cell,* **11**, 513-25.

Wang, J., Friedman, G., Doyon, Y., Wang, N. S., Li, C. J., Miller, J. C., Hua, K. L., Yan, J. J., Babiarz, J. E., Gregory, P. D. & Holmes, M. C. (2012). Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. *Genome Res.,* **22**, 1316-26.

Wang, X., Emelyanov, A., Korzh, V. & Gong, Z. (2003). Zebrafish atonal homologue zath3 is expressed during neurogenesis in embryonic development. *Dev. Dyn.,* **227**, 587-92.

Wardle, F., Odom, D., Bell, G., Yuan, B., Danford, T., Wiellette, E., Herbolsheimer, E., Sive, H., Young, R. & Smith, J. (2006). Zebrafish promoter microarrays identify actively transcribed embryonic genes. *Genome Biol.,* **7**, R71-R71.

West, S. C. (2003). Molecular views of recombination proteins and their control. *Nat. Rev. Mol. Cell Biol.,* **4**, 435-45.

Westerfield, M. (1993). *The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Brachydanio rerio),* Eugene, University of Oregon Press.

Wilson, S. W., Brand, M. & Eisen, J. S. (2002). Patterning the zebrafish central nervous system. *Results Probl. Cell Differ.,* **40**, 181-215.

Wodarz, A. & Huttner, W. B. (2003). Asymmetric cell division during neurogenesis in Drosophila and vertebrates. *Mech. Dev.,* **120**, 1297-1309.

Wolffe, A. P. & Guschin, D. (2000). Review: chromatin structural features and targets that regulate transcription. *J. Struct. Biol.,* **129**, 102-22.

Wullimann, M. F. (2009). Secondary neurogenesis and telencephalic organization in zebrafish and mice: a brief review. *Integrative zoology,* **4**, 123-33.

Xu, C., Fan, Z. P., Müller, P., Fogley, R., Dibiase, A., Trompouki, E., Unternaehrer, J., Xiong, F., Torregroza, I., Evans, T., Megason, S. G., Daley, G. Q., Schier, A. F., Young, R. A. & Zon, L. I. (2012). Nanog-like Regulates Endoderm Formation through the Mxtx2-Nodal Pathway. *Dev. Cell,* **22**, 625-38.

Xu, Q., Alldus, G., Holder, N. & Wilkinson, D. G. (1995). Expression of truncated Sek-1 receptor tyrosine kinase disrupts the segmental restriction of gene expression in the Xenopus and zebrafish hindbrain. *Development,* **121**, 4005-16.

Yang, H., Wang, H., Shivalila, C. S., Cheng, A. W., Shi, L. & Jaenisch, R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell,* **154**, 1370-9.

Yao, Z., Fong, A. P., Cao, Y., Ruzzo, W. L., Gentleman, R. C. & Tapscott, S. J. (2013). Comparison of endogenous and overexpressed MyoD shows enhanced binding of physiologically bound sites. *Skeletal muscle,* **3**, 8-8.

Yeyati, P. L., Shaknovich, R., Boterashvili, S., Li, J., Ball, H. J., Waxman, S., Nason-Burchenal, K., Dmitrovsky, E., Zelent, A. & Licht, J. D. (1999). Leukemia translocation protein PLZF inhibits cell growth and expression of cyclin A. *Oncogene,* **18**, 925-34.

Zaade, D., Schmitz, J., Benke, E., Klare, S., Seidel, K., Kirsch, S., Goldin-Lang, P., Zollmann, F. S., Unger, T. & Funke-Kaiser, H. (2013). Distinct signal transduction pathways downstream of the (P)RR revealed by microarray and ChIP-chip analyses. *PLoS ONE,* **8**, e57674-e57674.

Zhang, T., Xiong, H., Kan, L. X., Zhang, C. K., Jiao, X. F., Fu, G., Zhang, Q. H., Lu, L., Tong, J. H., Gu, B. W., Yu, M., Liu, J. X., Licht, J., Waxman, S., Zelent, A., Chen, E. &

Chen, S. J. (1999). Genomic sequence, structural organization, molecular evolution, and aberrant rearrangement of promyelocytic leukemia zinc finger gene. *Proc. Natl. Acad. Sci. U. S. A.,* **96**, 11422-7.

Zhang, X., Guo, C., Chen, Y., Shulha, H. P., Schnetz, M. P., Laframboise, T., Bartels, C. F., Markowitz, S., Weng, Z., Scacheri, P. C. & Wang, Z. (2008a). Epitope tagging of endogenous proteins for genome-wide ChIP-chip studies. *Nat. Methods,* **5**, 163-5.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W. & Liu, X. S. (2008b). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.,* **9**, R137-R137.

Zlatanova, J. & Thakar, A. (2008). H2A.Z: view from the top. *Structure,* **16**, 166-79.

Zu, Y., Tong, X., Wang, Z., Liu, D., Pan, R., Li, Z., Hu, Y., Luo, Z., Huang, P., Wu, Q., Zhu, Z., Zhang, B. & Lin, S. (2013). TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat. Methods*, 1-4.